(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 July 2003 (03.07.2003)

PCT

(10) International Publication Number
WO 03/054162 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/41014

(22) Internatinnal Filing Date:
19 December 2002 (19.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/029,397 20 December 2001 (20.12.2001) US

(71) Applicant *(for all designated States except US)*: **AMBION, INC.** [US/US]; 2130 Woodward St., Suite 200, Austin, TX 78744-1832 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MURPHY, George, L.** [US/US]; 8825 Escabosa Dr., Austin, TX 78748 (US). **WHITLEY, J., Penn** [US/US]; 5203 Avenue F, Austin, TX 78751 (US).

(74) Agent: **SHISHIMA, Gina, N.**; Fulbright & Jaworski L.L.P., Suite 2400, 600 Congress Avenue, Austin, TX 78701 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS

(57) Abstract: The present invention concerns a system for isolating, depleting, or separating a targeted nucleic acid, such as rRNA, from a sample comprising targeted and nontargeted nucleic acids. It effects a way of enriching for nontargeted nucleic acids, such as mRNAs. The invention further concerns methods of implementing the system and kits for implementing the system, which involves at least one bridging nucleic acid comprising 1) a targeting region complementary to a region on the targeted nucleic acid and 2) a bridging region complementary to the capture region of a capture nucleic acid that comprises a nonreactant structure. The nonreactant structure can be used to isolate the hybridizing molecules after incubation under conditions that allows hybridization.

WO 03/054162 A2

DESCRIPTION

METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates generally to the fields of molecular biology and microbial pathogenesis. More particularly, it concerns methods, compositions, and kits for isolating, depleting, separating a targeted nucleic acid population from other nucleic acid populations as a means for enriching those other nucleic acid population(s). More particularly, it concerns methods, compositions, and kits for enriching mRNA populations by depleting eukaryotic and/or prokaryotic rRNA from a sample using engineered bridging and capture nucleic acid molecules.

2. Description of Related Art

The ongoing efforts in microbial genome sequencing will enable unprecedented advances in our understanding of microbes and host-microbe interactions. Dozens of prokaryotic genomes, including those of numerous human pathogens, have been completely sequenced, and many others are in progress. Consequently, a renewal of focus and energy has emerged in the fields of microbial evolution, microbial pathogenesis, and infectious diseases. The potential impact of genomics on these disciplines is the subject of several recent reviews (Cummings *et al.*, 2000; Cornelis *et al.*, 2001; Fox *et al.*, 2001; *Current Opinion in Microbiology*). For host-microbe interactions, the ability to measure the expression of every single gene in a microorganism will make possible studies of such complex interactions as the global regulation of virulence factors and the mechanisms of response to host cells and their microenvironment. Scientists will also be able to evaluate the complete repertoire of host cell gene expression in response to the pathogen. Undoubtedly, novel interactions and responses between microbes and their hosts will be discovered, leading to a more complete picture of infectious diseases and how to control them.

In the past decade, researchers studying bacteria developed several novel approaches to evaluate global gene transcription in response to environmental stimuli, including host-microbe interactions. Prior to the era of genome sequencing, Chuang *et al.* (Chuang *et al.*, 1993) used an ordered set of *E. coli* lambda library clones to evaluate global transcription responses of *E. coli*. Other groups employed subtractive hybridization and differential screening to evaluate induction

of gene expression in *Mycobacterium avium* after phagocytosis by macrophages (Plum *et al.*, 1994) or in *Pyrococcus* grown under specific environmental conditions (Robinson *et al.*, 1994). Researchers further developed this approach with an elegant procedure for the selective capture of transcribed sequences (SCOTS) (Graham *et al.*, 1999). At the same time, many scientists bypassed library construction altogether and used using differential display (Liang *et al.*, 1995) to discover genes that are transcribed differently under various growth conditions. Although useful in certain circumstances, differential display is frequently a hit-or-miss prospect and gives no information on global transcription. More recently, serial analysis of gene expression (SAGE) (Velculescu *et al.*, 1995) emerged as a method for analyzing the complete transcriptome of a cell. SAGE, like differential display, can be useful but requires large amounts of nucleic acid sequencing. Not unexpectedly, for organisms whose genomes have been sequenced, array analysis is emerging as the method of choice for global gene expression studies with bacteria. Macroarrays (filter-based arrays) and microarrays (slide-based arrays) of complete genomes have made possible the simultaneous expression analysis of thousands of genes. The advent of microarray technology has already enabled analyses of the host response to interactions with pathogenic organisms (Cummings *et al.*, 2000). Similarly, microarray analysis and other methods have been used to evaluate gene expression in bacteria grown under different environmental conditions *in vitro*.

The application of array analysis to gene expression profiling in prokaryotes was an immediate outgrowth of similar studies with eukaryotic organisms, occurring only within the past two to three years. Infectious disease researchers have already begun applying microarray analysis to the study of complex host-microbe interactions. To date, such analyses of host-microbe interactions have been limited to the evaluation of host cell responses to bacteria or viruses. *Bordetella pertussis*, *Listeria monocytogenes*, *Neisseria meningitidis*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Salmonella dublin*, and *Staphylococcus aureus* are among the bacterial pathogens whose effects on host cell gene expression have been evaluated with microarrays. Array analyses of eukaryotic host cell transcription are feasible because of the ability to isolate polyadenylated mRNAs from eukaryotic cells and to specifically label mRNAs by oligo dT-primed cDNA synthesis.

Although it has been alluded to in the literature (Cummings *et al.*, 2000; Rappuoli, 2000), complete genome array expression analyses of bacteria in response to interactions with host cells

have not been widely published, if at all. Studies that examine the global bacterial gene response in the presence of host cells will require the development of tools to enable the efficient isolation, enrichment, and labeling of bacterial mRNAs (Cummings *et al.*, 2000; Graham *et al.*, 1999; Gingeras *et al.*, 2000; Graham *et al.*, 2001).

However, technical limitations of current methods available for purification and evaluation of bacterial mRNAs preclude these types of whole genome analysis. To realize the full potential of the genomics revolution, methods for purifying mRNAs from total bacterial RNA populations and particularly from mixtures of host cell and bacterial RNA need to be developed.

Isolating sufficient quantities of high quality bacterial mRNA is perhaps the most demanding technical requirement impeding analyses of bacterial gene expression in the presence of host cells. A small percentage of bacterial mRNAs may be A-tailed, but these are targeted for degradation and tend to be unstable. As a result, the commonly used method for mRNA purification with eukaryotic cells, oligo-dT capture, is ineffective.

Only a few studies have described methods for enriching or purifying bacterial mRNAs. Several groups (Plum *et al.*, 1994; Robinson *et al.*, 1994; Su *et al.*, 1998) have used rRNA subtraction to enrich for bacterial mRNAs. These procedures involved hybridization of rRNAs to biotinylated plasmid containing rRNA genes or to biotinylated antisense rRNAs followed by streptavidin capture and removal. This yields some benefits, but it requires fairly large amounts of plasmids or antisense RNA. Biotinylation of large amounts of DNA or RNA is often tricky and can be prohibitively expensive if biotin-modified nucleotides are incorporated during antisense RNA synthesis. In general, these methods have not seen widespread use. As mentioned above, Graham and Clarke-Curtiss (Graham *et al.*, 1999) went further in enriching for mycobacterial mRNAs with SCOTS. The SCOTS procedure is effective for detecting genes specifically expressed in the presence of host cells but is hampered by being a multi-step procedure that requires production of normalized double-stranded cDNA, PCR, differential hybridization, and cDNA capture. In addition to these methods, researchers have developed methods to polyadenylate bacterial mRNAs, thereby allowing for their purification by oligo dT-capture. Amara and Vijaya (Amara *et al.*, 1997) demonstrated that mRNAs in purified polysomes can be specifically polyadenylated and purified by oligo-dT capture. Wendisch et al. (Wendisch *et al.*, 2001) showed that the same process can be carried out with crude cell extracts.

Several shortcomings are associated with the polyadenylation approach. Different mRNAs may be polyadenylated to different extents or not at all depending on the structure of their 5' and 3' ends (Feng *et al.*, 2000). Polyadenylation in a cell lysate, followed by purification of RNA, will require inactivation of cellular RNAses so that transcripts are not degraded during the polyadenylation reaction. Optimizing the reaction to work reproducibly in many different bacterial cell lysates would likely be very difficult. Despite many worthy attempts, simple and universal procedures for bacterial mRNA enrichment, especially in the presence of host cell RNA, remain elusive. Thus, there remains a continued need for improvements in mRNA enrichment and/or the depletion of other RNA populations.

SUMMARY OF THE INVENTION

The present invention involves a system that allows for the isolation, separation, and depletion of a population of nucleic acid molecules. The system involves components that may be used to implement methods for isolating, separating, or depleting a targeted nucleic acid. Such components may also be included in kits of the invention.

In embodiments of the invention, a population of nucleic acids may be targeted for isolation, separation, or depletion. Such a nucleic acid is referred to as "targeted nucleic acid" or "targeted nucleic acid molecule." Alternatively, it may be referred to as a "nucleic acid target." In particular embodiments of the invention, the targeted nucleic acid is rRNA. In alternative embodiments, the targeted nucleic acid is mRNA, tRNA, or DNA including, cDNA and genomic DNA. The targeted nucleic acid may be in a sample, which is a composition that is suspected of containing the targeted nucleic acid. In some embodiments, the sample is obtained from or includes prokaryotes or eukaryotes or both. The sample may be cells, tissues, organs, and lysates, fractionations, or portions thereof. Furthermore, the targeted nucleic acid is targeted via a "targeting region" in the targeted nucleic acid. A "targeted region" refers to a region of the targeted nucleic acid that is complementary with the targeting region of a bridging nucleic acid and that allows the targeted nucleic acid to be separated from other non-targeted nucleic acid populations.

In embodiments in which the targeted nucleic acid is rRNA, the rRNA may be the 5S, 16S, or 23S rRNA from prokaryotes, though it may be any rRNA species from a prokaryotes. It is specifically contemplated that nucleic acids may be targeted in Gram positive bacteria and Gram negative bacteria. In further embodiments, the targeted rRNA is 5.8S, 17S or 18S, or 28S

rRNA (referred to as "types of rRNA") from a eukaryote. It is further contemplated that tRNA may be a targeted nucleic acid population either by itself or in combination with any of the targeted nucleic acids described herein. A non-limiting list of targeted rRNAs from various organisms is provided in a later section and is contemplated to be part of the invention.

In embodiments of the invention, the system involves a bridging nucleic acid, a capture nucleic acid, and a targeted nucleic acid, as shown, for example, in FIG. 1. While in many embodiments of the invention it is contemplated that the bridging nucleic acid and the capture nucleic acid are oligonucleotides, it is specifically contemplated that they may be polynucleotides as well. Thus, any embodiment involving an oligonucleotide may be implemented with a polynucleotide. Bridging nucleic acids, capture nucleic acids, and targeted nucleic acids of the invention may include, be at least or be at most 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 or more residues in length.

Furthermore, a "bridging nucleic acid" is a nucleic acid molecule that comprises a bridging region and a targeting region, while a "capture nucleic acid" is a nucleic acid molecule that comprises a capture region. It is contemplated that bridging, targeting, and capture regions of the invention may be, be at least or be at most 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830,

840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 residues in length.

A "bridging nucleic acid" refers to a molecule that includes nucleic acid residues or analogs and that includes at least one targeting region and at least one bridging region. A "targeting region" refers to a region of the molecule that is involved in targeting a particular nucleic acid or nucleic acid population and is thus complementary to all or part of the sequence of the targeted nucleic acid. It is further contemplated that more than one targeting region may be included in a bridging nucleic acid. The bridging nucleic acid may include or have up to 2, 3, 4, 5, 6, 7, 8, 9, 10, or more targeting regions. When there are multiple targeting regions, it is contemplated that the regions may be complementary to different, nonoverlapping sequences from the same targeted nucleic acid or they may be complementary to similar or overlapping sequences from the same targeted nucleic acid, or they may be complementary to sequences in different targeted nucleic acids. While mRNA may be targeted, it is specifically contemplated that mRNA is not targeted and thus the targeting region does not have a stretch of polypyrimidine residues, such as poly-T or poly-U to hybridize to the poly-A tail of eukaryotic mRNA. Also considered part of the invention is using single or multiple bridging nucleic acids to deplete an rRNA population. In some embodiments, a single bridging nucleic acid may contain one or more targeting regions that are complementary to different types of rRNA ("types" refer to sizes based on intact lengths). Thus, in some embodiments, the largest type of rRNA may be targeted ("largest" refers to longest nucleic acid molecule when intact, even though molecules that are no longer intact may also be targeted if they retain the sequence that is complementary to all or part of a targeting region). In still further embodiments, the second largest rRNA or the first and second largest rRNA types may be targeted by a single bridging nucleic acid with targeting regions to each or to more than one nucleic acid, each with a targeting region to a different type of rRNA. In still further embodiments, a bridging nucleic acid has a targeting region complementary to one or more of the following prokaryotic and eukaryotic rRNA types: 5S, 16S, 23S, 5.8S, 17S, 18S, and/or 28S. A bridging nucleic acid may target 1, 2, 3, 4, 5, 6, 7, or more types of rRNA, as well as any and all tRNA types, both eukaryotic and prokaryotic.

A "bridging region" in a bridging nucleic acid refers to a region that mediates an interaction with a capture nucleic acid. In further embodiments, the bridging region is a

polypurine or polypyrimidine stretch of residues. A bridging region can include a stretch of adenine or guanine residues or cytosine, uracil, or thymidine residues. In some embodiments, it is contemplated that more than one bridging region is included in a bridging nucleic acid, such as 2, 3, 4, 5, or more bridging regions.

A "capture nucleic acid" refers to a molecule that includes nucleotides or nucleotide analogs, a capture region, and a nonreacting structure. A "capture region" refers to a region that interacts with the bridging region of a bridging nucleic acid. In embodiments of the invention, the bridging region and the capture region are complementary to each other and hybridize to one another under conditions that allow for hybridization of complementary regions. In some embodiments of the invention, the capture region and bridging region are a stretch of complementary repeated nucleotides (complementary homopolymeric regions). For example, they may be homopolymeric A, T, G, C, or U. In other embodiments of the invention, however, the bridging and capture regions are any sequence, so long as they are complementary. In some embodiments of the invention, the capture region has a sequence that includes at least 5, 6, 7, 8, 9, 10 or more contiguous nucleotides of SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, and SEQ ID NO:92 (collectively referred to as "SEQ ID NOs: 87-92"). In some embodiments, the capture regions comprises any of of SEQ ID NOs 87-92.

There may be more than one nonreacting structure attached, covalently or noncovalently, to a capture nucleic acid. There may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nonreacting structures as part of a capture nucleic acid.

A capture nucleic acid also includes a "nonreacting structure," which refers to a compound that does not chemically react with a nucleic acid. In some embodiments, a nonreacting structure is a magnetic bead or rod, which allows the capture nucleic acid, a bridging nucleic acid and a target nucleic acid to be isolated from a sample with a magnetic field, such as a magnetic stand. In still further embodiments, the nonreacting structure is a bead or other structure that can be physically captured, such as by using a basket, filter, or by centrifugation. It is contemplated that a bead may include plastic, glass, teflon, silica, a magnet or be magnetizeable, a metal such as a ferrous metal or gold, carbon, cellulose, latex, polystyrene, and other synthetic polymers, nylon, cellulose, nitrocellulose, polymethacrylate, polyvinylchloride, styrene-divinylbenzene, or any chemically-modified plastic or any other nonreacting structure. In still further embodiments, the nonreacting structure is biotin or iminobiotin. Biotin or

iminobiotin binds to avidin or streptavidin , which can be used to isolate the capture nucleic acid and any hybridizing molecules. Furthermore, in some embodiments of the invention, the nonreacting structure is cellulose or an analog thereof.

It is contemplated that the location of the targeting and bridging regions in the bridging nucleic acid may be at a variety of positions. The location of targeted regions in a targeted nucleic acid or a capture region in a capture nucleic acid may also vary. The location of any of these regions or nonreacting structure may be or be within 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 12, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000 or more nucleotides from the 3' and/or 5' end of the relevant nucleic acid ("relevant nucleic acid" refers to the nucleic acid in which the region is located). Moreover, it is contemplated that a region, such as a bridging, capture, targeted, or targeting region—as well as a nonreacting structure—may be at or within 100-5000 residues, 150-4000 residues, 200-3000 residues, 250-2000 residues, 300-1500 residues, 350-1000 residues, 400-900 residues, 450-800 residues, or 500-700 residues of the 5' or 3' end of the relevant nucleic acid.

Furthermore, it is also contemplated that the spacing between regions may vary. Regions in the same nucleic acid or a region and a nonreacting structure may be adjacent to one another or there may be residues between them or between each of them. The number of intervening residues may be the following or may be at least or at most of the following: 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590,

600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, or more nucleotides between them or each of them.

As for the location of the sequence to which the targeting region is complementary, termed "targeted region," this may be 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000 nucleotides or more from the 3' and/or 5' end of the targeted nucleic acid. It is specifically contemplated that the targeting region hybridizes to a sequence located between 100 and 5000, 150 and 4000, 200 and 3000, 250 and 2000, and 300 and 1000 residues of the 5' and/or 3' end of the targeted nucleic acid. It is also contemplated that the targeted region is at the 3' or 5' end of the targeted nucleic acid. Alternatively, the targeted region may not be within 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000 or more nucleotides from the termini of a targeted nucleic acid.

In some embodiments, the targeting region comprises or is complementary to all or 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940,

950, 960, 970, 980, 990, 1000 or more contiguous nucleotides of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:72, or SEQ ID NO:73 (collectively referred to as "SEQ ID NOS:1-73"), as well as SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:91, and SEQ ID NO:92 (collectively referred to as "SEQ ID NOS:1-92"). It is specifically contemplated that targeting regions of the invention comprise, in some embodiments, at least 5 contiguous nucleotides of SEQ ID NO:1-22; it is also contemplated that targeting regions of the invention are complementary to a sequence ("sequence" in the context of complementary regions refers to a sequence of at least 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, or more nucleotides in length) of SEQ ID NOS:23-86, which are sequences of rRNA molecules.

It will be understood that any embodiment discussed with respect to nucleotides applies also when nucleotide analogs are used. It is specifically contemplated that nucleotide analogs may be employed with respect to bridging and capture nucleic acids of the invention.

It is contemplated that nucleic acids of the invention include RNA, DNA, locked nucleic acid™ (LNA), iso-bases, and/or peptide mimetics. It is contemplated that all or part of nucleic acids of the invention may include such nucleic acid components.

The present invention further concerns methods of isolating and/or depleting nucleic acids from a sample. In some embodiments, methods include a) incubating a sample with a first

bridging nucleic acid comprising (1) at least one bridging region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the first targeting region and the targeted nucleic acid; b) incubating the first bridging nucleic acid with a capture nucleic acid comprising a nonreacting structure and a capture region comprising at least 5 nucleic acid residues, under conditions that allowing hybridization between the first bridging region and the capture region. In additional embodiments, one or more other steps may be included in combination with the method discussed above. Other steps involve isolating the targeted nucleic acid from the remainder of the sample; discarding the portion of the sample that hybridizes directly or indirectly to the capture nucleic acid (indirect hybridization refers to specific association of compounds that occurs through hybridization with a mediating compound, for example, indirect hybridization of a capture nucleic acid and a targeted nucleic acid via hybridization to a bridging nucleic acid); incubating the sample with additional bridging nucleic acids, under conditions allowing hybridization between the targeting region of the additional bridging nucleic acid and the targeted nucleic acid; implementing the method with respect to other targeted nucleic acids; washing the capture nucleic acid after incubation with the sample and the bridging nucleic acid; incubating the capture nucleic acid, bridging nucleic acid, and sample with elution buffer after isolating the targeted nucleic acid from the rest of the sample; eluting the targeted nucleic acid from the nonreactant structure; using the capture nucleic acid in a subsequent method involving a new sample; discarding the targeted nucleic acid after separating it from the sample; performing hybridizations between the bridging nucleic acid and the sample and the capture nucleic acid and the sample at the same temperatures or at different temperatures; performing the above hybridization steps at the same time, sequentially (one after the other or the other after the one); exposing the sample to a magnetic field or magnet, particularly when a magnetic bead or other object comprises all or part of the non-reacting structure of the capture nucleic acid; and incubating the sample with streptavidin or avidin, particularly if biotin or iminobiotin is used as a non-reacting structure.

In some embodiments of the invention, the sample, a bridging nucleic acid and/or a capture nucleic acid are incubated in a buffer, which, in some embodiments, includes TEAC or TMAC.

In methods of the invention involving more than one bridging nucleic acid, it is contemplated that the targeting region of the first bridging nucleic acid may be complementary to

a different sequence of a different targeted nucleic acid than a targeting region of another bridging nucleic acid. Alternatively, different bridging nucleic acids may have targeting regions that are complementary to the same targeted nucleic acid. In the latter case, it is further contemplated that the targeting regions be complementary to sequences that overlap one another or ma be complementary to sequences in non-overlapping locations.

In cases in which targeting regions are complementary to different targeted nucleic acids, embodiments may involve targeting the largest rRNA molecule in a sample with one bridging nucleic acid and the second largest rRNA molecule in a sample with another bridging nucleic acid. In still further embodiments, another or third bridging nucleic acid will target the third largest rRNA molecule in a sample, while another or a fourth bridging nucleic acid will target the fourth largest rRNA molecule in a sample.

In another embodiment of the invention, there is a method for depleting rRNA from a sample comprising incubating the sample with (1) at least a first bridging oligonucleotide comprising a bridging region comprising a polypurine region of at least 5 residues in length and a targeting region comprising at least 5 contiguous residues complementary to an rRNA molecule in the sample and (2) a capture oligonucleotide comprising a magnetic bead and a capture region comprising a polypyrimidine region of at least 5 residues in length, under conditions allowing hybridization between the bridging oligonucleotide and the capture oligonucleotide and between the bridging oligonucleotide and the rRNA; b) incubating the sample with a magnetic bead; and c) isolating the magnetic bead. In still further embodiments, the first bridging oligonucleotide comprises a targeting region complementary to prokaryotic 23S rRNA. In still further embodiments, there is a second bridging oligonucleotide with a targeting region complementary to a different region of a prokaryotic 23S RNA than the first bridging oligonucleotide. In even further embodiments, there is a third and fourth bridging oligonucleotide each with a targeting region complementary to different sequences of a prokaryotic 16S rRNA.

As discussed earlier, a sample may be depleted or isolated as a way of enriching for the nontargeted nucleic acid, such as mRNA. In further embodiments of the invention, enriched mRNA can be used to prepare cDNA according to methods known to those of ordinary skill in the art, and as described herein. Thus, in cases in which mRNA is enriched as a result of methods of the invention, embodiments may further include discarding the portion of the sample

that hybridizes to the capture oligonucleotide. More specifically targeted rRNA may be discarded and the mRNA remaining in the sample may be used to produce cDNA molecules. cDNA molecules may be used in a variety of methods, including, but not limited to, library production, production of proteins, and for creating and screening arrays. Therefore, in some embodiments of the invention, cDNA made from mRNA enriched according to methods of the invention are attached to a solid support or surface so as to create a nucleic acid array. The term "nucleic acid array" refers to a plurality of target elements, wherein each target element comprising one or more nucleic acid molecules immobilized on one or more solid surfaces at discrete locations to which sample nucleic acids can by hybridized. The nonreacting solid surface or support may be any of a number of materials, including plastic, glass, or nylon. In some embodiments, the solid support is a plate. The plate may have wells that contain the target elements. Plates may have 2, 3, 4, 5, 6, 7, 8, 9, 10 or more wells ("multi-well"), and up to at least 96 or 192 wells. In some embodiments of the invention, the sample nucleic acids comprise cDNAs made by depleting a sample of rRNA, according to methods of the invention. Those embodiments may further involve contacting a nucleic acid array with the cDNA. Alternatively, cDNA made according to the invention may be used as target elements on an array. In any of these embodiments of the invention, it is specifically contemplated that enriched mRNA may be amplified into RNA or DNA by techniques known to those of skill in the art and then used in methods of the invention, such as to probe or screen an array.

The present invention also concerns kits that include compositions of the invention to implement the methods discussed herein. These kits can be used for the depletion, isolation, or purification of nucleic acids. Kits contain these compositions in a suitable container means.

In some embodiments, a kit includes 1) at least one capture oligonucleotide comprising a capture region and a magnetic bead; and 2) at least a first bridging oligonucleotide comprising i) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and ii) at least one targeting region comprising 10 contiguous nucleic acids complementary to an rRNA.

In additional embodiments, there is a second bridging oligonucleotide comprising i) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and ii) at least one targeting region comprising 10 contiguous nucleic acids complementary to an rRNA. In some kits, the targeting region of the second bridging

oligonucleotide is complementary to the same rRNA as the targeting region of the first bridging oligonucleotide, while in other embodiments, these are complementary to different rRNAs. Further embodiments involve kits in which the targeting region of the first bridging oligonucleotide is complementary to the largest rRNA of a prokaryote or eukaryote. In other embodiments, the second bridging oligonucleotide has a targeting region that is complementary to either the largest rRNA of a prokaryote or eukaryote or the second largest rRNA of a prokaryote or eukaryote. It is specifically contemplated that kits may include one or more bridging oligonucleotides targeting prokaryotic rRNA (16S, 23S, or both) *and* one or more bridging oligonucleotides targeting eukaryotic rRNA (18S, 28S, or both); thus, a kit may be used for depleting both eukaryotic and prokaryotic rRNA, in some embodiments.

Kits may also include a third, fourth, fifth, sixth, seventh, eighth, ninth, tenth or more bridging oligonucleotides with targeting region complementary to the same or different rRNAs as the targeting regions of the first and second bridging oligonucleotides. It is contemplated that the targeting regions of the bridging oligonucleotides in kits of the invention may be complementary to prokaryote 16S rRNA, prokaryote 23S rRNA, prokaryote 5S rRNA, eukaryote 17S or 18S rRNA, eukaryote 28SrRNA, and/or eukaryote 5.8S rRNA. It is further contemplated that targeting regions of bridging oligonucleotides in kits may have all or part of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, or SEQ ID NO:22 (collectively referred to as "SEQ ID NOS:1-22"). Alternatively, kits may include targeting regions as discussed above with respect to SEQ ID NOS:23-86, *i.e.* targeting regions complementary to a sequence from SEQ ID NOS:23-86. Kits of the invention may also include one or more of the following: binding buffer with TMAC, binding buffer with TEAC, magnetic stand, wash solution, nuclease-free water; RNAse inhibitors, glycogen, control RNA, sodium acetate, ammonium acetate, streptavidin beads, avidin beads, magnetic beads, beads of any nonreacting structure--including those discussed above--capture basket; capture filters, RNA markers, nuclease-free containers such as tubes and tips, and any other composition described herein.

It is contemplated that kits of the invention may be used to implement methods of the invention, that methods of the invention may be implemented with compositions of the invention, and that kits may include any composition of the invention.

It is further contemplated that kits, methods, and compositions of the invention may effect a depletion of a targeted nucleic acid in a sample by reducing its amount in the sample by at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 or more percent.

Kits of the invention also include materials for creating a nucleic acid array and probing a nucleic acid array. Any of the kits discussed above may also include a solid support for preparing a nucleic acid array.

The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one." When the term "at least" is used in the context of bridging, targeting, or capture regions, as well as for capture and bridging oligonucleotides, it is contemplated that there is an upper limit of 20 for practical purposes, even though more such regions or oligonucleotides could be implemented with the invention. Furthermore, it should be understood that a number (cardinal or ordinal) used in the context of compositions of the invention refers to a "kind" of that composition; thus, "a first oligonucleotide" in the context of a "second oligonucleotide" refers to "one of that kind of oligonucleotide," and not one single oligonucleotide molecule.

Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

FIG. 1. Depiction of molecules in system. A bridging oligonucleotide is shown with a targeting region and a bridging region. The targeting region is complementary to a targeted region in the targeted nucleic acid, which is an rRNA molecule. The bridging region is complementary to the capture region in the capture oligonucleotide, which is attached, by way of example, to a magnetic bead as a nonreacting structure.

FIG. 2A-1 to A-14 and FIG. 2B-1 to B-27. Sequence comparison of different rRNAs from different bacteria to *E. coli* rRNA with MegAlign sequence analysis software version 4.05 from DNA Star, Incorporated. **A.** The 5' end of the sequence is shown on the first page of the figure in FIG. 2A-1 and continues until the last page of the figure, FIG. 2A-14, in which the 3' end of the same sequence is shown. Shown is a sequence comparison of 16S rRNA of listed prokaryotic organisms to 16S rRNA from *E. coli* (SEQ ID NO:34). The sequences are the 16S rRNA from the following organisms: B. subtilis (SEQ ID NO:23); B. anthracis (SEQ ID NO. 24); E. faecalis (SEQ ID NO. 25); L. lactis (SEQ ID NO. 26); L. monocyt (SEQ ID NO. 27); S. aureus (SEQ ID NO. 28); S. mutans (SEQ ID NO. 29); S. pneumon (SEQ ID NO. 30); S. pyogenes (SEQ ID NO. 31); M. avian (SEQ ID NO. 32); M. tuberculosis (SEQ ID NO. 33); K. pneumoniae (SEQ ID NO. 35); A. actino (SEQ ID NO. 36); H. influenzae (SEQ ID NO. 37); E. bronchiseptica (SEQ ID NO. 38); B. parapertussis (SEQ ID NO. 39); B. pertussis (SEQ ID NO. 40); B. cepacia (SEQ ID NO. 41); B. mallei (SEQ ID NO. 42); B. pseudomallei (SEQ ID NO. 43); N. gonorrhoeae (SEQ ID NO. 44); N. mening (SEQ ID NO. 45); P. aeruginosa (SEQ ID NO. 46); V. cholerae (SEQ ID NO. 47); and Y. enterocolitica (SEQ ID NO. 48). **B.** The 5' end of the sequence is shown on the first page of the figure in FIG. 2B-1 and continues until the last page of the figure, FIG. 2B-27, in which the 3' end of the same sequence is shown. Shown is a sequence comparison of 23S rRNA of listed prokaryotic organisms to 23S rRNA from *E. coli* (SEQ ID NO:60). The sequences are the 23S rRNA from the following organisms: B. subtilis (SEQ ID NO:49); B. anthracis (SEQ ID NO. 50); E. facaelis (SEQ ID NO. 51); L. lactis (SEQ ID

NO. 52); L. monocytogenes (SEQ ID NO. 53); S. aureus (SEQ ID NO. 54); S. mutans (SEQ ID NO. 55); S. pneumoniae (SEQ ID NO. 56); S. pyogenes (SEQ ID NO. 57); M. avium (SEQ ID NO. 58); M. tuberculosis (SEQ ID NO. 59); K. pneumoniae (SEQ ID NO. 61); H. influenzae (SEQ ID NO. 62); B. bronchiseptica (SEQ ID NO. 63); B. parapertussis (SEQ ID NO. 64); B. pertussis (SEQ ID NO. 65); B. cepacia (SEQ ID NO. 66); E. mallei (SEQ ID NO. 67); E. pseudomallei (SEQ ID NO. 68); N. gonorrhoeae (SEQ ID NO. 69); N. eminigititdis (SEQ ID NO. 70); P. aeruginosa (SEQ ID NO. 71); V. cholerae (SEQ ID NO. 72); Y. enterocolitica (SEQ ID NO. 73).

FIG. 3. Electropherograms of RNA from a control reaction. *E. coli* total RNA was purified with RNAwiz™ (Ambion) and carried through the rRNA depletion procedure as described in Example 2, except that bridging nucleic acids were left out of the reaction. A sample of the RNA was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 4. Electropherograms of RNA from an experimental reaction after ribosomal RNA depletion. *E. coli* total RNA was purified with RNAwiz™ (Ambion) and carried through the rRNA depletion procedure as described in Example 2. A sample of the RNA was analyzed as described in the legend to FIG. 3.

FIG. 5A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Example 5, but with no bridging oligonucleotides. The sample contains *E. coli* and rat liver total RNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 5 with bridging oligonucleotides. The sample is depleted of *E. coli* 16S and 23S rRNA and rat liver 18S and 28S rRNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 6A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Example 6, but with no bridging oligonucleotides. The sample contains human liver total RNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 6 with bridging oligonucleotides. The sample is depleted of human 18S and 28S rRNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 7A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Example 7, but with no bridging oligonucleotides. The sample contains rat liver total RNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 6 with bridging oligonucleotides. The sample is depleted of rat 18S and 28S rRNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 8A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Example 6, but with no bridging oligonucleotides. The sample contains mouse liver total RNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 8 with bridging oligonucleotides. The sample is depleted of mouse 18S and 28S

rRNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 9A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Example 11 with no bridging oligonucleotides. The sample contains human total RNA (50 μg) and *E. coli* total RNA (500 ng). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 11 with bridging oligonucleotides. The sample is depleted of human 18S and 28S rRNA, but *E. coli* total RNA remains in the sample. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 10A-C. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction with no bridging oligonucleotides performed as described in Example 12. The sample contains rat liver total RNA (25 μg) and *E. coli* total RNA (2 μg). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 12 with bridging oligonucleotides. The sample is depleted of rat 18S and 28S rRNA, but *E. coli* total RNA remains in the sample. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **C.** Agilent

2100 Bioanalyzer electropherogram of a sample from an experimental reaction performed as described in Example 12 with bridging oligonucleotides. The sample is depleted of rat 18S and 28S rRNA and *E. coli* 16S and 23S rRNA. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 11A-B. *E. coli* gene arrays probed with cDNA from experiments. Experimental reactions were performed as described in Example 13. RNA samples were used to generate radiolabeled cDNA for use as probes with replicate portions of Sigma-Genosys Panorama™ *E. coli* gene arrays. **A.** *E. coli* gene array probed with a sample from a control reaction. The sample contains human total RNA (25 μg) and *E. coli* total RNA (2 μg). **B.** *E. coli* gene array probed with an RNA sample that was depleted of human 18S and 28S rRNA and *E. coli* 16S and 23S rRNA.

FIG. 12A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Examples 1 and 2 with no bridging oligonucleotides. The sample contains *Campylobacter fetus* total RNA (10 μg). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction with 10 μg of *Campylobacter fetus* total RNA that employed the bridging oligonucleotides d16S-807, d16S-1092, d23S-479CH, and d23S-2511. The sample is depleted of 16S rRNA, 23S rRNA fragment (1260 nt), and 23S rRNA fragment (1667 nt). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 13A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Examples 1 and 2 with no bridging oligonucleotides. The sample contains *Rhodobacter sphaeroides* total RNA (10 μg). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper

Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction with 10 μg of *Rhodobacter sphaeroides* total RNA that employed the bridging oligonucleotides d16S-537 (16 pmol), d16S-1114R(16 pmol), d23S-479CH (16 pmol), d23S-1954 (16 pmol), and d23S-2511 (16 pmol). The sample is depleted of the 16S rRNA and the 23S rRNA fragment that co-migrates with the 16S rRNA. The sample is also depleted of the 23S rRNA fragment (1260 nt). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

FIG. 14A-B. Electropherograms of RNA from experiments. **A.** Agilent 2100 Bioanalyzer electropherogram of a sample from a control reaction performed as described in Examples 1 and 2 with no bridging oligonucleotides. The sample contains *Anabaena sp.* total RNA (10 μg). The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01). **B.** Agilent 2100 Bioanalyzer electropherogram of a sample from an experimental reaction with 10 μg of *Anabaena sp.* total RNA that employed the bridging oligonucleotides d16S-364 (12.5 pmol), d16S-1087CY (12.5 pmol), d23S-485 (20 pmol), and d23S-1954 (35 pmol). The sample is depleted of 16S rRNA and the 23S rRNA fragments at 520 nt, 2090 nt, and 2470 nt. The RNA sample was analyzed with the RNA 6000 Lab Chip Kit® (Caliper Technologies Corp.) using the Agilent 2100 Bioanalyzer (Agilent Technologies). The electropherogram shown was generated with Agilent 2100 Bioanalyzer Bio Sizing software (Version A.02.01).

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention concerns a system for isolating, depleting, or identifying specific, targeted nucleic acid populations, such as rRNA in a sample, in some cases for the purpose of enriching for other nucleic acid populations. The targeted nucleic acid, components of the system, and the methods for implementing the system, as well as variations thereof, are provided below.

I. **Targeted Nucleic Acid**

The present invention concerns targeting a particular nucleic acid population (*i.e.*, mRNA, rRNA, tRNA, genomic DNA) or targeting types of a nucleic acid population, such as individual tRNAs, rRNAs (5S, 16S, or 23S rRNA from prokaryotes; 5.8S, 17S or 18S, or 28S from eukaryotes), or specific mRNAs. A nucleic acid is targeted by using a bridging nucleic acid that has a targeting region—a region complementary to all or part of the targeted nucleic acid.

In some embodiments, the invention is specifically concerned with depleting or isolating rRNA from other nucleic acids ("non-targeted nucleic acids" or "enriched population"). The 5S, 16S, and/or 23S rRNA from a prokaryote may be the targeted nucleic acid. Also, the 5.8S, 17S (observed in yeast) or 18S, and/or 28S from a eukaryote may be the targeted nucleic acid. Alternatively, rRNAs from both prokaryotes and eukaryotes may be targeted, such as with a sample that has eukaryotic host cells infected with a prokaryotic organism. The sequences for ribosomal RNAs are well known to those or ordinary skill in the art and can be readily found in sequence databases such as GenBank (www.ncbi.nlm.nih.gov/) or are published. Nucleic acids may be targeted by targeting regions that are complementary to all or part of the targeted nucleic acid. Targeted nucleic acids may be, be at least, or be at most 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5100, 5200, 5300, 5400, 5500, 5600, 5700, 5800, 5900, 6000, 6100, 6200, 6300, 6400, 6500, 6600, 6700, 6800, 6900, 7000, 7100, 7200, 7300, 7400, 7500, 7600, 7700, 7800, 7900, 8000, 8100, 8200, 8300, 8400, 8500, 8600, 8700, 8800, 8900, 9000, 9100, 9200, 9300, 9400, 9500, 9600, 9700, 9800, 9900, 10000, or more nucleotides in length. Furthermore, any region of at least five contiguous nucleotides in the targeted nucleic acid may be used as the targeted region—that is, the region that is complementary to the targeting region of a bridging nucleic acid. Also, there may be more than one targeted region in a targeted nucleic acid. There may be, be at least, or be at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or

more targeted regions in a targeted nucleic acid. A targeted region may be a region in a targeted nucleic acid that has greater than 70%, 80%, or 90% homology with a sequence from a different targeted nucleic acid. In some embodiments, the targeted region from a targeted nucleic acid is identical to a sequence in a different targeted nucleic acid. For example, 23S rRNA of various prokaryotes may be targeted using a targeted region common to a group of organisms, such as Gram negative bacteria or a subset of such bacteria. Alternatively, a targeted region may be a sequence unique to a particular targeted nucleic acid. However, for purposes of this application, a "targeted region" is not a poly-A region, such as a poly-A tail of an eukaryotic mRNA. Additional information regarding targeted rRNAs is provided below. This information is provided as an example of targeted nucleic acids. However, it is contemplated that there may be sequence variations from individual organism to organism and these sequences provided as simply an example of one sequenced nucleic acid, even though such variations exist in nature. It is contemplated that these variations may also be targeted, and this may or may not require changes to a targeting nucleic acid or to the hybridization conditions, depending on the variation, which one of ordinary skill in the art could evaluate and determine.

A number of patents concern a targeted nucleic acid, for example, U.S. Patent Nos. 4,486,539; 4,563,419; 4,751,177; 4,868,105; 5,200,314; 5,273,882; 5,288,609; 5,457,025; 5,500,356; 5,589,335; 5,702,896; 5,714,324; 5,723,597; 5,759,777; 5,897,783; 6,013,440; 6,060,246; 6,090,548; 6,110,678; 6,203,978; 6,221,581; 6,228,580; and WO 01/32672, all of which are specifically incorporated herein by reference.

A. Prokaryotic rRNA

Prokaryotic rRNA can be a targeted nucleic acid of the invention. The following examples are provided, but the invention is not limited solely to these organisms and sequences (GenBank accession number provided and/or region within sequence that corresponds to the targeted rRNA):

1. **Superkingdom Archaea** (archaebacteria)

Aeropyrum pernix	
16S	D83259
Aeropyrum pernix NC_000854	
APErRNA05 (16S)	1218763-1220185
APErRNA03 (23S)	1213627-1218039
Methanococcus jannaschii	
16S	M59126

Methanococcus jannaschii NC_000909
MJrrnA16S 157985-159459
MJrrnA23S 154759-157648
Halobacterium marismortui
23S X13738
Halobacterium sp. NRC-1 NC_002607
rrs (16S) 1875505-1876977
rrlA (23S) 1877506-1880411
Thermoplasma acidophilum
23S M32298
Thermoplasma acidophilum NC_002578
16S 1475300-1475770

2. Superkingdom Eubacteria (eubacteria)

a. Firmicutes (Gram-positive bacteria)

i) Bacillus/Clostridium group (low G+C gram-positive bacteria)

Listeria innocua Clip 11262 NC_003212
16S 260527-262081
23S 262327-265257

Listeria monocytogenes strain EGD NC_003210
16S 237466-239020
23S 239265-242195

Bacillus subtilis NC000964
RrnO 16S 9809-11361
RrnA 23S 11707-14634

Bacillus anthracis
16S (1508nt) AF155950
23S (2922nt) AF267877

Bacillus thuringiensis
16S (1486nt) D16281
23S (2923nt) AF267880

Staphylococcus aureus strain Mu50 NC_002758
16S 530479-532033
23S 532398-535231

Staphylococcus aureus N315 NC_002745
SarRNA01 16S 506138-507692
SarRNA02 23S 508166-510999

Clostridium acetobutylicum ATCC824 NC_003030
16SarRNA 9710-11219
23SarRNA 11398-14303

Clostridium difficile
16S (1470nt) X73450

Clostridium perfringens
16S M69264 (499-2294)

Mycoplasma genitalium G37 L43967
MgrnA16S 170009-171527
MgrnA23S 171730-174463

Mycoplasma pneumoniae NC_000912
16S 118312-119824
23S 120057-122961

Mycoplasma pulmonis NC_002771
16S 813583-815113
23S 810563-813297

Streptococcus pneumoniae R6 NC_003098
RRNA16S-1 15161-16674
RRNA23S-1 16945-19846

Streptococcus pneumoniae TIGR4 AE005672
SprnaA16S 15394-16806
SprnaA23S 17142-20043

Streptococcus pyogenes AE004092
16S 17170-18504
23S 19037-21937

Streptococcus mutans
16S (1334nt) X58303
23S AF139599 (1940-4840)

Lactococcus lactis
16S X64887 (508-2055)
23S X64887 (2360-5257)

Enterococcus faecalis
16S (1449nt) Y18293
23S (2912nt) AJ295306

ii) **Actinobacteria (high G+C gram-positive bacteria)**
Mycobacterium leprae strain TN NC_002677
Rrs16S 1341144-1342692
Rrl23S 1342976-1346100

Mycobacterium tuberculosis CDC 1551 NC_002755
MtrnaA16S 1471388-1472923
MtrnaA23S 1473199-1476336

Mycobacterium avium
16S (1372nt) M61673
23S X74494 (295-3401)

Corynebacterium glutamicum
16S (1479nt) Z46753

Rhodococcus equi
16S (1478nt) X80614

b. **Spirochaetales (spirochetes)**

Borrelia burgdorferi AE000783
RrlB 16S 444581-446118
RrlB 23S 438590-441508

Treponema pallidum AE000520
TprmaA16S 230162-231656
TprmaA23S 231950-234850

Borrelia burgdorferi
16S AE001147 (9459-10996)
23S AE001147 (212-3145)

c. **Thermotogales**

Thermotoga maritima AE000512
TmrmaA16S 188968-190526
TmrmaA23S 190766-193787

d. **Thermus/Deinococcus group**

Deinococcus radiodurans R1 NC_001263
DrrmaA16S 2285518-2287019
DrrmaA23S 2245319-2246194

Deinococcus radiodurans
16S AE002076 (7275-8776)
23S AE001886 (8829-10771)

e. **Chlamydiales (chlamydias)**

Chlamydia trachomatis AE001273
16SrRNA1 854128-855677
23SrRNA1 855993-858862

Chlamydophila pneumoniae AR39 NC_002179
CprmA16S 1069329-1070785
CprmA23S 1066159-1069022

Chlamydophila psittaci
16S U68447 (1-1553)
23S U68447 (1778-4721)

f. **Proteobacteria (purple bacteria)**

i) **Alpha subdivision**
Rickettsia conorii Malish 7 NC_003103

Rrs16S 884601-886108
Rrl23S 281797-284557

Rickettsia prowazekii strain Madrid E AJ235269
Rrs16S 772263-773769
Rrl23S 257853-260613

Rickettsia typhi
16S (1444nt) M20499
23S Y13133 (956-3716)

Ehrlichia bovis
16S (1488nt) U03775

Agrobacterium tumefaciens C58 AE007870
16S 768991-770427
23S 765313-767565

Brucella melitensis
16S AF220148 (645-2129)
23S AF220148 (2896-3024..3204-5807)

Rhizobium rhizogenes
16S (1369nt) D13945

ii) **Beta subdivision**

Neisseria meningitides strain MC58 AE002098
NmrmaA16S 60971-62514
NmrmaA23S 63178-66068

Bordetella bronchiseptica
16S (1532nt) X57026
23S (2865nt) X70371

Bordetella parapertussis
16S (1464nt) U04949
23S (2865nt) X68368

Bordetella pertussis
16S (1464nt) U04950

Burkholderi mallei
16S (1488nt) AF110188
23S (2882nt) Y17183

Burkholderi pseudomallei
16S (1488nt) U91839
23S (2882nt) Y17184

Neisseria gonorrhoeae
16S (1544nt) X07714
23S (2890nt) X67293

iii) **Gamma group**

Buchnera sp. APS NC_002528

Rrs 16S	274065-275524
Rrl 23S	539539-542451

Escherichia coli K12 U00096

RrsH 16S	223771-225312
RrlH 23S	225759-228662

Escherichia coli 0157:H7 NC_002695

RrsH 16S	227102-228643
RrlH 23S	229090-231992

Salmonella enterica serovar Typhi NC_003198

16S	287479-289020
23S	289375-292380

Salmonella typhimurium LT2 NC_003197

RrsH16S	289189-290732
RrlH23S	291244-294336

Yersinia pestis NC_003143

16S	12292-13763
23S	14272-17178

Klebsiella pneumoniae

16S (1534nt)	X87276
23S (2903nt)	X87284

Yersinia enterocolitica

16S (1484nt)	Z49830
23S (2906nt)	U77925

Proteus vulgaris

16S (2067nt)	X07652

Shigella flexneri

16S (1468nt)	X80679

Shigella sonnei

16S (1467nt)	X80726

Shigella dysenterica

16S (1487nt)	X96966

Haemophilus influenzae Rd L42023

HirmE16S	1511137-1512634
HirmE23S	123801-126697

Pasteurella multocida

16S (1543nt)	M35018

Actinobacillus actinomycetemcomitans
16S (1485nt) M75037

Actinobacillus pleuropneumoniae
16S D30032 (83-1625)

Haemophilus somnus
16S (1483nt) M75046

Legionella pneumophila
16S (1544nt) M59157

Mannheimia haemolytica
16S (1472nt) U57072

Vibrio cholerae chromosomel NC_002505
16Sa rRNA 53823-55357
23Sa rRNA 55784-58670

Vibrio parahaemolyticus
16S (1499nt) M59161

Coxiella burnetii
16S (1484nt) M21291
23S X79704 (1620-3350)

Aeromonas hydrophila
16S (1538nt) X87271

Aeromonas salmonicida
16S (1502nt) X60405

Francisella tularesis
16S (1517nt) Z21931

Moraxella catarrhalis
16S (1511nt) U10876

Pseudomonas aeruginosa AE004091
16S 722096-726631
23S 724103-726993

Pseudomonas putida
16S (1527nt) D84020

iv) Delta/Epsilon subdivisions

Campylobacter jejuni AL111168
16S 39249-40761
23S 41568-44457

Helicobacter pylori 26695 NC_000915
HPrmB16S 1511137-1512634
HPrmB23S 1473918-1476893

g. Cyanobacteria

Synechocystis sp. PCC 6803 NC_000911
Rrn16Sa 2452187-2453675
Rrn23Sa 2448839-2451721

Synechococcus sp. (Anacystis nidulans)
16S X03538 (1432-2918)
23S X00512 (251-3126)

h. CFB/Green sulfer bacteria group

Porphyromonas gingivalis
16S (1474nt) L16492

B. Eukaryotic rRNA

Targeted nucleic acids of the invention may also be one or more types of eukaryotic rRNAs. Eukaryotes include, but are not limited to: mammals, fish, birds, amphibians, fungi, and plants. The following provides sequences for some of these targeted nucleic acids. It is contemplated that other eukaryotic rRNA sequences can be readily obtained by one of ordinary skill in the art, and thus, the invention includes, but is not limited to, the sequences shown below.

Superkingdom Eukaryota (eucaryotes)

Homo sapiens (human)
18S M10098
18S K03432
18S X03205
28S M11167

Mus muculus
18S X00686
28S X00525

Rattus norvegicus
18S M11188
18S X01117

Rattus norvegicus V01270.1
18S 1-1874
28S 3862-8647

II. Isolation and/or Depletion System Nucleic Acids

The present invention concerns compositions comprising a nucleic acid or a nucleic acid analog in a system or kit to deplete, isolate, or separate a nucleic acid population from other nucleic acid populations, for which enrichment may be desirable. It concerns a bridging nucleic acid and a capture nucleic acid to deplete, isolate, or separate out a targeted nucleic acid, as discussed above.

A. Bridging Nucleic Acids

Bridging nucleic acids of the invention comprise a bridging region and a targeting region. As discussed in other sections, the location of these regions may be throughout the molecule, which may be of a variety of lengths. The bridging nucleic acid may comprise RNA, DNA, both, or analogs of either or both.

The bridging region comprises a sequence that is complementary to at least five contiguous nucleotides in the capture nucleic acid. It is contemplated that that this region may be a homogenous sequence, that is, have the same nucleotide repeated across its length, such as a repeat of A, C, G, T, or U residues. However, to avoid hybridizing with a poly-A tailed mRNA in a sample comprising eukaryotic nucleic acids, it is contemplated that most embodiments will not have a poly-U or poly-T bridging region when dealing with such samples having poly-A tailed RNA. In some embodiments, the bridging region is a poly-C region and the capture region is a poly-G region, or vice versa. In other embodiments, the bridging region will be a random sequence that is complementary to the capture region (or the capture region will be random and the bridging region will be complementary to it). In further embodiments, the bridging region will have a designed sequence that is not homopolymeric but that is complementary to the capture region or vice versa. Sequences may be determined empirically. In many embodiments, it is preferred that this will be a random sequence or a defined sequence that is not a homopolymer. Some sequences will be determined empirically during evaluation in the assay.

B. Capture Nucleic Acids

Capture nucleic acids of the invention comprise a capture region and a nonreacting structure that allows the capture nucleic acid, any molecules specifically binding or hybridizing to the capture nucleic acid—such as the bridging nucleic acid—and any molecules specifically binding or hybridizing to the bridging nucleic acid—such as the targeted nucleic acid—to be isolated away from other nucleic acid populations.

The capture nucleic acid may comprise RNA, DNA, both, or analogs of either or both. However, in some embodiments of the invention, it is specifically contemplated to be homopolymeric (only one type of nucleotide residue in molecule, such as poly-C), though in other embodiments, it is specifically contemplated not to be homopolymeric and be heteropolymeric, as described for bridging regions.

1. Capture Regions

The main requirement for bridging and capture nucleic acid sequences is that they are complementary to one another. The capture region may be a poly-pyrimidine or poly-purine region comprising at least 5 nucleic acid residues. In addition, it may be heteropolymeric, either a random sequence or a designed sequence that is complementary to the bridging region of the nucleic acid with which it should hybridize.

In addition to the capture oligos already described herein, the following are also considered for use with the present invention:

NRS-5'-TAACCTGGTCGTAAC-3' (SEQ ID NO:87)

NRS-5'-CCCCCCCCCCCCCC-3' (SEQ ID NO:88)

NRS-5'-GCCCCTAACCTCGTCG (SEQ ID NO:89)

NRS-5'-CGGCCCTAGCCGGGTCGTACCTCCGG (SEQ ID NO:90)

NRS-5'-CGGCCCTAACCTGGTCGTAACTCCGG (SEQ ID NO:91)

NRS-5'-AGGCTTCGATCCCGGGATCCGCG (SEQ ID NO:92)

As discussed below, "NRS" refers to a non-reacting structure.

2. Nonreacting Structures (NRS)

A nonreacting structure is a compound or structure that will not react chemically with nucleic acids, and in some embodiments, with any molecule that may be in a sample. Nonreacting structures may comprise plastic, glass, teflon, silica, a magnet, a metal such as gold, carbon, cellulose, latex, polystyrene, and other synthetic polymers, nylon, cellulose, nitrocellulose, polymethacrylate, polyvinylchloride, styrene-divinylbenzene, or any chemically-

modified plastic. They may also be porous or non-porous materials. The structure may also be a particle of any shape that allows the targeted nucleic acid to be isolated, depleted, or separated. It may be a sphere, such as a bead, or a rod, or a flat-shaped structure, such as a plate with wells. Also, it is contemplated that the structure may be isolated by physical means or electromagnetic means. For example, a magnetic field may be used to attract a non-reacting structure that includes a magnet. The magnetic field may be in a stand or it may simply be placed on the side of a tube with the sample and a capture nucleic acid that is magnetized. Examples of physical ways to separate nucleic acids with their specifically hybridizing compounds are well known to those of skill in the art. A basket or other filter means may be employed to separate the capture nucleic acid and its hybridizing compounds (direct and indirect). The non-reacting structure and sample with nucleic acids of the invention may be centrifuged, filtered, dialyzed, or captured (with a magnet). When the structure is centrifuged it may be pelleted or passed through a centrifugible filter apparatus. The structure may also be filtered, including filtration using a pressure-driven system. Many such structures are available commercially and may be utilized herewith. Other examples can be found in WO 86/05815, WO90/06045, U.S. Patent 5,945,525, all of which are specifically incorporated by reference.

Cellulose is a structural polymer derived from vascular plants. Chemically, it is a linear polymer of the monosaccharide glucose, using β, 1-4 linkages. Cellulose can be provided commercially, including from the Whatman company, and can be chemically sheared or chemically modified to create preparations of a more fibrous or particulate nature. CF-1 cellulose from Whatman is an example that can be implemented in the present invention.

Synthetic plastic or glass beads may be employed in the context of the invention. The beads may be complexed with avidin or streptavidin and they may also be magnetized. The complexed streptavidin can be used to capture biotin linked to an oligo-dT or -U or poly (dT) or poly(U) moiety, either before or after hybridization to the poly(A) tails of mRNA. Alternatively, the oligo/poly(dT/U) moiety can be attached to the beads directly through chemical coupling. The beads may be collected using gravity- or pressure-based systems and/or filtration devices. If the beads are magnetized, a magnet can be used to separate the beads from the rest of the sample. The magnet may be employed with a stand or a stick or other type of physical structure to facilitate isolation.

Other components include isolation apparatuses such as filtration devices, including spin filters or spin columns.

C. Nucleic Acid Compositions

Embodiments of the present invention concern bridging, capture, and targeted nucleic acids. In particular aspects, a targeted nucleic acid encodes for or comprises a transcribed nucleic acid. In other aspects, a bridging nucleic acid comprises a targeting region that comprises a nucleic acid segment having the sequence of all or part of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:72, or SEQ ID NO:73 (collectively referred to as "SEQ ID NOS:1-73"), as well as SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, and SEQ ID NO:92 (collectively referred to as "SEQ ID NOS:1-92"). In particular aspects, a targeted nucleic acid encodes a protein, polypeptide, peptide. Nucleic acids of the invention comprise RNA, DNA, analogs of RNA, analogs of DNA, or a combination thereof.

The term "nucleic acid" is well known in the art. A "nucleic acid" as used herein will generally refer to a molecule (*i.e.*, a strand) of DNA, RNA or a derivative or analog thereof, comprising a nucleobase. A nucleobase includes, for example, a naturally occurring purine or pyrimidine base found in DNA (*e.g.*, an adenine "A," a guanine "G," a thymine "T" or a cytosine "C") or RNA (*e.g.*, an A, a G, an uracil "U" or a C). The term "nucleic acid" encompass the

terms "oligonucleotide" and "polynucleotide," each as a subgenus of the term "nucleic acid." The term "oligonucleotide" refers to a molecule of between about 3 and about 100 nucleobases in length. The term "polynucleotide" refers to at least one molecule of greater than about 100 nucleobases in length.

These definitions generally refer to a single-stranded molecule, but in specific embodiments will also encompass an additional strand that is partially, substantially or fully complementary to the single-stranded molecule. Thus, a nucleic acid may encompass a double-stranded molecule or a triple-stranded molecule that comprises one or more complementary strand(s) or "complement(s)" of a particular sequence comprising a molecule. As used herein, a single stranded nucleic acid may be denoted by the prefix "ss," a double stranded nucleic acid by the prefix "ds," and a triple stranded nucleic acid by the prefix "ts."

1. Nucleobases

As used herein a "nucleobase" refers to a heterocyclic base, such as for example a naturally occurring nucleobase (*i.e.*, an A, T, G, C or U) found in at least one naturally occurring nucleic acid (*i.e.*, DNA and RNA), and naturally or non-naturally occurring derivative(s) and analogs of such a nucleobase. A nucleobase generally can form one or more hydrogen bonds ("anneal" or "hybridize") with at least one naturally occurring nucleobase in manner that may substitute for naturally occurring nucleobase pairing (*e.g.*, the hydrogen bonding between A and T, G and C, and A and U).

"Purine" and/or "pyrimidine" nucleobase(s) encompass naturally occurring purine and/or pyrimidine nucleobases and also derivative(s) and analog(s) thereof, including but not limited to, those of a purine or pyrimidine substituted by one or more of an alkyl, caboxyalkyl, amino, hydroxyl, halogen (*i.e.*, fluoro, chloro, bromo, or iodo), thiol or alkylthiol moiety. Preferred alkyl (*e.g.*, alkyl, caboxyalkyl, etc.) moieties comprise of from about 1, about 2, about 3, about 4, about 5, to about 6 carbon atoms. Other non-limiting examples of a purine or pyrimidine include a deazapurine, a 2,6-diaminopurine, a 5-fluorouracil, a xanthine, a hypoxanthine, a 8-bromoguanine, a 8-chloroguanine, a bromothymine, a 8-aminoguanine, a 8-hydroxyguanine, a 8-methylguanine, a 8-thioguanine, an azaguanine, a 2-aminopurine, a 5-ethylcytosine, a 5-methylcyosine, a 5-bromouracil, a 5-ethyluracil, a 5-iodouracil, a 5-chlorouracil, a 5-propyluracil, a thiouracil, a 2-methyladenine, a methylthioadenine, a N,N-diemethyladenine, an azaadenines, a 8-bromoadenine, a 8-hydroxyadenine, a 6-hydroxyaminopurine, a 6-thiopurine, a

4-(6-aminohexyl/cytosine), and the like. A table of non-limiting, purine and pyrimidine derivatives and analogs is also provided herein below.

Table 1-Purine and Pyrimidine Derivatives or Analogs

Abbr.	Modified base description	Abbr.	Modified base description
ac4c	4-acetylcytidine	Mam5s2u	5-methoxyaminomethyl-2-thiouridine
Chm5u	5-(carboxyhydroxylmethyl) uridine	Man q	Beta,D-mannosylqueosine
Cm	2'-O-methylcytidine	Mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine
Cmnm5s2u	5-carboxymethylamino-methyl-2-thioridine	Mcm5u	5-methoxycarbonylmethyluridine
Cmnm5u	5-carboxymethylaminomethyluridine	Mo5u	5-methoxyuridine
D	Dihydrouridine	Ms2i6a	2-methylthio-N6-isopentenyladenosine
Fm	2'-O-methylpseudouridine	Ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl)carbamoyl)threonine
Gal q	Beta,D-galactosylqueosine	Mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methyl-carbamoyl)threonine
Gm	2'-O-methylguanosine	Mv	Uridine-5-oxyacetic acid methylester
I	Inosine	o5u	Uridine-5-oxyacetic acid (v)
I6a	N6-isopentenyladenosine	Osyw	Wybutoxosine
m1a	1-methyladenosine	P	Pseudouridine

Table 1-Purine and Pyrimidine Derivatives or Analogs			
Abbr.	Modified base description	Abbr.	Modified base description
m1f	1-methylpseudouridine	Q	Queosine
m1g	1-methylguanosine	s2c	2-thiocytidine
m1I	1-methylinosine	s2t	5-methyl-2-thiouridine
m22g	2,2-dimethylguanosine	s2u	2-thiouridine
m2a	2-methyladenosine	s4u	4-thiouridine
m2g	2-methylguanosine	T	5-methyluridine
m3c	3-methylcytidine	t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)carbamoyl)threonine
m5c	5-methylcytidine	Tm	2'-O-methyl-5-methyluridine
m6a	N6-methyladenosine	Um	2'-O-methyluridine
m7g	7-methylguanosine	Yw	Wybutosine
Mam5u	5-methylaminomethyluridine	X	3-(3-amino-3-carboxypropyl)uridine, (acp3)u

A nucleobase may be comprised of a nucleoside or nucleotide, using any chemical or natural synthesis method described herein or known to one of ordinary skill in the art.

2. Nucleosides

As used herein, a "nucleoside" refers to an individual chemical unit comprising a nucleobase covalently attached to a nucleobase linker moiety. A non-limiting example of a "nucleobase linker moiety" is a sugar comprising 5-carbon atoms (*i.e.*, a "5-carbon sugar"), including but not limited to a deoxyribose, a ribose, an arabinose, or a derivative or an analog of a 5-carbon sugar. Non-limiting examples of a derivative or an analog of a 5-carbon sugar include a 2'-fluoro-2'-deoxyribose or a carbocyclic sugar where a carbon is substituted for an oxygen atom in the sugar ring.

Different types of covalent attachment(s) of a nucleobase to a nucleobase linker moiety are known in the art. By way of non-limiting example, a nucleoside comprising a purine (*i.e.*, A

or G) or a 7-deazapurine nucleobase typically covalently attaches the 9 position of a purine or a 7-deazapurine to the 1'-position of a 5-carbon sugar. In another non-limiting example, a nucleoside comprising a pyrimidine nucleobase (*i.e.*, C, T or U) typically covalently attaches a 1 position of a pyrimidine to a 1'-position of a 5-carbon sugar.

3. Nucleotides

As used herein, a "nucleotide" refers to a nucleoside further comprising a "backbone moiety". A backbone moiety generally covalently attaches a nucleotide to another molecule comprising a nucleotide, or to another nucleotide to form a nucleic acid. The "backbone moiety" in naturally occurring nucleotides typically comprises a phosphorus moiety, which is covalently attached to a 5-carbon sugar. The attachment of the backbone moiety typically occurs at either the 3'- or 5'-position of the 5-carbon sugar. However, other types of attachments are known in the art, particularly when a nucleotide comprises derivatives or analogs of a naturally occurring 5-carbon sugar or phosphorus moiety.

4. Nucleic Acid Analogs

A nucleic acid may comprise, or be composed entirely of, a derivative or analog of a nucleobase, a nucleobase linker moiety and/or backbone moiety that may be present in a naturally occurring nucleic acid. As used herein a "derivative" refers to a chemically modified or altered form of a naturally occurring molecule, while the terms "mimic" or "analog" refer to a molecule that may or may not structurally resemble a naturally occurring molecule or moiety, but possesses similar functions. As used herein, a "moiety" generally refers to a smaller chemical or molecular component of a larger chemical or molecular structure. Nucleobase, nucleoside and nucleotide analogs or derivatives are well known in the art, and have been described (see for example, Scheit, 1980, incorporated herein by reference).

Additional non-limiting examples of nucleosides, nucleotides or nucleic acids comprising 5-carbon sugar and/or backbone moiety derivatives or analogs, include those in U.S. Patent No. 5,681,947 which describes oligonucleotides comprising purine derivatives that form triple helixes with and/or prevent expression of dsDNA; U.S. Patents 5,652,099 and 5,763,167 which describe nucleic acids incorporating fluorescent analogs of nucleosides found in DNA or RNA, particularly for use as fluorescent nucleic acids probes; U.S. Patent 5,614,617 which describes oligonucleotide analogs with substitutions on pyrimidine rings that possess enhanced nuclease stability; U.S. Patents 5,670,663, 5,872,232 and 5,859,221 which describe oligonucleotide

analogs with modified 5-carbon sugars (*i.e.*, modified 2'-deoxyfuranosyl moieties) used in nucleic acid detection; U.S. Patent 5,446,137 which describes oligonucleotides comprising at least one 5-carbon sugar moiety substituted at the 4' position with a substituent other than hydrogen that can be used in hybridization assays; U.S. Patent 5,886,165 which describes oligonucleotides with both deoxyribonucleotides with 3'-5' internucleotide linkages and ribonucleotides with 2'-5' internucleotide linkages; U.S. Patent 5,714,606 which describes a modified internucleotide linkage wherein a 3'-position oxygen of the internucleotide linkage is replaced by a carbon to enhance the nuclease resistance of nucleic acids; U.S. Patent 5,672,697 which describes oligonucleotides containing one or more 5' methylene phosphonate internucleotide linkages that enhance nuclease resistance; U.S. Patents 5,466,786 and 5,792,847 which describe the linkage of a substituent moiety, which may comprise a drug or label to the 2' carbon of an oligonucleotide to provide enhanced nuclease stability and ability to deliver drugs or detection moieties; U.S. Patent 5,223,618 which describes oligonucleotide analogs with a 2 or 3 carbon backbone linkage attaching the 4' position and 3' position of adjacent 5-carbon sugar moiety to enhanced cellular uptake, resistance to nucleases and hybridization to target RNA; U.S. Patent 5,470,967 which describes oligonucleotides comprising at least one sulfamate or sulfamide internucleotide linkage that are useful as nucleic acid hybridization probe; U.S. Patents 5,378,825, 5,777,092, 5,623,070, 5,610,289 and 5,602,240 which describe oligonucleotides with three or four atom linker moiety replacing phosphodiester backbone moiety used for improved nuclease resistance, cellular uptake and regulating RNA expression; U.S. Patent 5,858,988 which describes hydrophobic carrier agent attached to the 2'-O position of oligonucleotides to enhanced their membrane permeability and stability; U.S. Patent 5,214,136, which describes oligonucleotides conjugated to anthraquinone at the 5' terminus that possess enhanced hybridization to DNA or RNA; enhanced stability to nucleases; U.S. Patent 5,700,922 which describes PNA-DNA-PNA chimeras wherein the DNA comprises 2'-deoxy-erythro-pentofuranosyl nucleotides for enhanced nuclease resistance, binding affinity, and ability to activate RNase H; and U.S. Patent 5,708,154 which describes RNA linked to a DNA to form a DNA-RNA hybrid. Other analogs that may be used with compositions of the invention include U.S. Patent 5,216,141 (discussing oligonucleotide analogs containing sulfur linkages), U.S. Patent 5,432,272 (concerning oligonucleotides having nucleotides with heterocyclic bases), and U.S. Patents 6,001,983, 6,037,120, 6,140,496 (involving oligonucleotides with non-standard bases), all of which are incorporated by reference.

5. Polyether and Peptide Nucleic Acids and Locked Nucleic Acids

In certain embodiments, it is contemplated that a nucleic acid comprising a derivative or analog of a nucleoside or nucleotide may be used in the methods and compositions of the invention. A non-limiting example is a "polyether nucleic acid", described in U.S. Patent Serial No. 5,908,845, incorporated herein by reference. In a polyether nucleic acid, one or more nucleobases are linked to chiral carbon atoms in a polyether backbone.

Another non-limiting example is a "peptide nucleic acid", also known as a "PNA", "peptide-based nucleic acid analog" or "PENAM", described in U.S. Patent Serial Nos. 5,786,461, 5891,625, 5,773,571, 5,766,855, 5,736,336, 5,719,262, 5,714,331, 5,539,082, and WO 92/20702, each of which is incorporated herein by reference. Peptide nucleic acids generally have enhanced sequence specificity, binding properties, and resistance to enzymatic degradation in comparison to molecules such as DNA and RNA (Egholm *et al.*, 1993; PCT/EP/01219). A peptide nucleic acid generally comprises one or more nucleotides or nucleosides that comprise a nucleobase moiety, a nucleobase linker moiety that is not a 5-carbon sugar, and/or a backbone moiety that is not a phosphate backbone moiety. Examples of nucleobase linker moieties described for PNAs include aza nitrogen atoms, amido and/or ureido tethers (see for example, U.S. Patent No. 5,539,082). Examples of backbone moieties described for PNAs include an aminoethylglycine, polyamide, polyethyl, polythioamide, polysulfinamide or polysulfonamide backbone moiety.

In certain embodiments, a nucleic acid analogue such as a peptide nucleic acid may be used to inhibit nucleic acid amplification, such as in PCR, to reduce false positives and discriminate between single base mutants, as described in U.S. Patent Serial No. 5,891,625. Other modifications and uses of nucleic acid analogs are known in the art, and are encompassed by the bridging and capture nucleic acids of the invention. In a non-limiting example, U.S. Patent 5,786,461 describes PNAs with amino acid side chains attached to the PNA backbone to enhance solubility of the molecule. In another example, the cellular uptake property of PNAs is increased by attachment of a lipophilic group. Several alkylamino moieties used to enhance cellular uptake of a PNA are described in U.S. Patent Nos. 5,766,855, 5,719,262, 5,714,331 and 5,736,336, which describe PNAs comprising naturally and non-naturally occurring nucleobases and alkylamine side chains that provide improvements in sequence specificity, solubility and/or binding affinity relative to a naturally occurring nucleic acid.

Another non-limiting example is a locked nucleic acid or "LNA." An LNA monomer is a bicyclic compound that is structurally similar to RNA nucleosides. LNAs have a furanose conformation that is restricted by a methylene linker that connects the 2'-O position to the 4'-C position, as described in Koshkin *et al*, 1998a and 1998b and Wahlestedt *et al.*, 2000.

6. Preparation of Nucleic Acids

A nucleic acid may be made by any technique known to one of ordinary skill in the art, such as for example, chemical synthesis, enzymatic production or biological production. Non-limiting examples of a synthetic nucleic acid (*e.g.*, a synthetic oligonucleotide), include a nucleic acid made by *in vitro* chemical synthesis using phosphotriester, phosphite or phosphoramidite chemistry and solid phase techniques such as described in EP 266,032, incorporated herein by reference, or via deoxynucleoside H-phosphonate intermediates as described by Froehler *et al.*, 1986 and U.S. Patent No. 5,705,629, each incorporated herein by reference. In the methods of the present invention, one or more oligonucleotide may be used. Various different mechanisms of oligonucleotide synthesis have been disclosed in for example, U.S. Patents. 4,659,774, 4,816,571, 5,141,813, 5,264,566, 4,959,463, 5,428,148, 5,554,744, 5,574,146, 5,602,244, each of which is incorporated herein by reference.

A non-limiting example of an enzymatically produced nucleic acid include one produced by enzymes in amplification reactions such as PCR™ (see for example, U.S. Patent 4,683,202 and U.S. Patent 4,682,195, each incorporated herein by reference), or the synthesis of an oligonucleotide described in U.S. Patent No. 5,645,897, incorporated herein by reference. A non-limiting example of a biologically produced nucleic acid includes a recombinant nucleic acid produced (*i.e.*, replicated) in a living cell, such as a recombinant DNA vector replicated in bacteria (see for example, Sambrook *et al.* 1989, incorporated herein by reference).

7. Purification of Nucleic Acids

A nucleic acid may be purified on polyacrylamide gels, cesium chloride centrifugation gradients, or by any other means known to one of ordinary skill in the art (see for example, Sambrook *et al.*, 1989, incorporated herein by reference).

In certain aspect, the present invention concerns a nucleic acid that is an isolated nucleic acid. As used herein, the term "isolated nucleic acid" refers to a nucleic acid molecule (e.g., an RNA or DNA molecule) that has been isolated free of, or is otherwise free of, the bulk of the

total genomic and transcribed nucleic acids of one or more cells. In certain embodiments, "isolated nucleic acid" refers to a nucleic acid that has been isolated free of, or is otherwise free of, bulk of cellular components or in vitro reaction components such as for example, macromolecules such as lipids or proteins, small biological molecules, and the like.

8. Nucleic Acid Segments

In certain embodiments, the nucleic acid comprises a nucleic acid segment. As used herein, the term "nucleic acid segment," are smaller fragments of a nucleic acid, such as for non-limiting example, those that correspond to targeted, targeting, bridging, and capture regions. Thus, a "nucleic acid segment" may comprise any part of a gene sequence, of from about 2 nucleotides to the full length of a targeted nucleic acid, capture nucleic acid, or bridging nucleic acid.

Various nucleic acid segments may be designed based on a particular nucleic acid sequence, and may be of any length. By assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all nucleic acid segments can be created:

n to n + y

where n is an integer from 1 to the last number of the sequence and y is the length of the nucleic acid segment minus one, where n + y does not exceed the last number of the sequence. Thus, for a 10-mer, the nucleic acid segments correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the nucleic acid segments correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the nucleic segments correspond to bases 1 to 20, 2 to 21, 3 to 22 ... and so on. In certain embodiments, the nucleic acid segment may be a probe or primer. As used herein, a "probe" generally refers to a nucleic acid used in a detection method or composition. As used herein, a "primer" generally refers to a nucleic acid used in an extension or amplification method or composition.

9. Nucleic Acid Complements

The present invention also encompasses a nucleic acid that is complementary to a other nucleic acids of the invention and targeted nucleic acids. More specifically, a targeting region in a bridging nucleic acid is complementary to the targeted region of the targeted nucleic acid and a bridging region of the bridging nucleic acid is complementary to a capture region of a capture nucleic acid. In particular embodiments the invention encompasses a nucleic acid or a nucleic

acid segment identical or complementary to all or part of the sequences set forth in SEQ ID NOS:1-92. A nucleic acid is "complement(s)" or is "complementary" to another nucleic acid when it is capable of base-pairing with another nucleic acid according to the standard Watson-Crick, Hoogsteen or reverse Hoogsteen binding complementarity rules. Unless otherwise specified, a nucleic acid region is "complementary" to another nucleic acid region if there is at least 70, 80%, 90% or 100% Watson-Crick base-pairing (A:T or A:U, C:G) between or between at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500 or more contiguous nucleic acid bases of the regions. As used herein "another nucleic acid" may refer to a separate molecule or a spatial separated sequence of the same molecule.

As used herein, the term "complementary" or "complement(s)" also refers to a nucleic acid comprising a sequence of consecutive nucleobases or semiconsecutive nucleobases (*e.g.*, one or more nucleobase moieties are not present in the molecule) capable of hybridizing to another nucleic acid strand or duplex even if less than all the nucleobases do not base pair with a counterpart nucleobase. In certain embodiments, a "complementary" nucleic acid comprises a sequence in which at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%, and any range derivable therein, of the nucleobase sequence is capable of base-pairing with a single or double stranded nucleic acid molecule during hybridization, as described in the Examples. In certain embodiments, the term "complementary" refers to a nucleic acid that may hybridize to another nucleic acid strand or duplex under conditions described in the Examples, as would be understood by one of ordinary skill in the art.

In certain embodiments, a "partly complementary" nucleic acid comprises a sequence that may hybridize in low stringency conditions to a single or double stranded nucleic acid, or contains a sequence in which less than about 70% of the nucleobase sequence is capable of base-pairing with a single or double stranded nucleic acid molecule during hybridization.

10. Hybridization

As used herein, "hybridization", "hybridizes" or "capable of hybridizing" is understood to mean the forming of a double or triple stranded molecule or a molecule with partial double or

triple stranded nature. The term "anneal" as used herein is synonymous with "hybridize." The term "hybridization", "hybridize(s)" or "capable of hybridizing" encompasses the terms "stringent condition(s)" or "high stringency" and the terms "low stringency" or "low stringency condition(s)."

As used herein "stringent condition(s)" or "high stringency" are those conditions that allow hybridization between or within one or more nucleic acid strand(s) containing complementary sequence(s), but precludes hybridization of random sequences. Stringent conditions tolerate little, if any, mismatch between a nucleic acid and a target strand. Such conditions are well known to those of ordinary skill in the art, and are preferred for applications requiring high selectivity. Non-limiting applications include isolating a nucleic acid, such as a gene or a nucleic acid segment thereof, or detecting at least one specific mRNA transcript or a nucleic acid segment thereof, and the like.

Stringent conditions may comprise low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.15 M NaCl at temperatures of about 50°C to about 70°C. Alternatively, stringent conditions may be determined largely by temperature in the presence of a TMAC solution with a defined molarity such as 3M TMAC. For example, in 3 M TMAC, stringent conditions include the following: for complementary nucleic acids with a length of 15 bp, a temperature of 45 °C to 55 °C; for complementary nucleotides with a length of 27 bases, a temperature of 65 °C to 75 °C; and, for complementary nucleotides with a length of >200 nucleotides, a temperature of 90 °C to 95°C. The publication of Wood *et al.*, 1985, which is specifically incorporated by reference, provides examples of these parameters. It is understood that the temperature and ionic strength of a desired stringency are determined in part by the length of the particular nucleic acid(s), the length and nucleobase content of the target sequence(s), the charge composition of the nucleic acid(s), and to the presence or concentration of formamide, tetramethylammonium chloride or other solvent(s) in a hybridization mixture.

It is also understood that these ranges, compositions and conditions for hybridization are mentioned by way of non-limiting examples only, and that the desired stringency for a particular hybridization reaction is often determined empirically by comparison to one or more positive or negative controls. Depending on the application envisioned it is preferred to employ varying conditions of hybridization to achieve varying degrees of selectivity of a nucleic acid towards a target sequence. In a non-limiting example, identification or isolation of a related target nucleic

acid that does not hybridize to a nucleic acid under stringent conditions may be achieved by hybridization at low temperature and/or high ionic strength. Such conditions are termed "low stringency" or "low stringency conditions", and non-limiting examples of low stringency include hybridization performed at about 0.15 M to about 0.9 M NaCl at a temperature range of about 20°C to about 50°C. Of course, it is within the skill of one in the art to further modify the low or high stringency conditions to suite a particular application.

11. Oligonucleotide Synthesis

Oligonucleotide synthesis is performed according to standard methods. See, for example, Itakura and Riggs (1980). Additionally, U.S. Patent 4,704,362; U.S. Patent 5,221,619, U.S. Patent . 5,583,013 each describe various methods of preparing synthetic structural genes.

Oligonucleotide synthesis is well known to those of skill in the art. Various different mechanisms of oligonucleotide synthesis have been disclosed in for example, U.S. Patents. 4,659,774, 4,816,571, 5,141,813, 5,264,566, 4,959,463, 5,428,148, 5,554,744, 5,574,146, 5,602,244, each of which is incorporated herein by reference.

Basically, chemical synthesis can be achieved by the diester method, the triester method polynucleotides phosphorylase method and by solid-phase chemistry. These methods are discussed in further detail below.

Diester method. The diester method was the first to be developed to a usable state, primarily by Khorana and co-workers. (Khorana, 1979). The basic step is the joining of two suitably protected deoxynucleotides to form a dideoxynucleotide containing a phosphodiester bond. The diester method is well established and has been used to synthesize DNA molecules (Khorana, 1979).

Triester method. The main difference between the diester and triester methods is the presence in the latter of an extra protecting group on the phosphate atoms of the reactants and products (Itakura *et al.*, 1975). The phosphate protecting group is usually a chlorophenyl group, which renders the nucleotides and polynucleotide intermediates soluble in organic solvents. Therefore purification's are done in chloroform solutions. Other improvements in the method include (i) the block coupling of trimers and larger oligomers, (ii) the extensive use of high-

performance liquid chromatography for the purification of both intermediate and final products, and (iii) solid-phase synthesis.

Polynucleotide phosphorylase method. This is an enzymatic method of DNA synthesis that can be used to synthesize many useful oligodeoxynucleotides (Gillam *et al.*, 1978; Gillam *et al.*, 1979). Under controlled conditions, polynucleotide phosphorylase adds predominantly a single nucleotide to a short oligodeoxynucleotide. Chromatographic purification allows the desired single adduct to be obtained. At least a trimer is required to start the procedure, and this primer must be obtained by some other method. The polynucleotide phosphorylase method works and has the advantage that the procedures involved are familiar to most biochemists.

Solid-phase methods. Drawing on the technology developed for the solid-phase synthesis of polypeptides, it has been possible to attach the initial nucleotide to solid support material and proceed with the stepwise addition of nucleotides. All mixing and washing steps are simplified, and the procedure becomes amenable to automation. These syntheses are now routinely carried out using automatic DNA synthesizers.

Phosphoramidite chemistry (Beaucage, and Lyer, 1992) has become by far the most widely used coupling chemistry for the synthesis of oligonucleotides. As is well known to those skilled in the art, phosphoramidite synthesis of oligonucleotides involves activation of nucleoside phosphoramidite monomer precursors by reaction with an activating agent to form activated intermediates, followed by sequential addition of the activated intermediates to the growing oligonucleotide chain (generally anchored at one end to a suitable solid support) to form the oligonucleotide product.

12. Expression Vectors

Other ways of creating nucleic acids of the invention include the use of a recombinant vector created through the application of recombinant nucleic acid technology known to those of skill in the art or as described herein. A recombinant vector may comprise a bridging or capture nucleic acid, particularly one that is a polynucleotide, as opposed to an oligonucleotide. An expression vector can be used create nucleic acids that are lengthy, for example, containing multiple targeting regions or relatively lengthy targeting regions, such as those greater than 100 residues in length.

The term "vector" is used to refer to a carrier nucleic acid molecule into which a nucleic acid sequence can be inserted for introduction into a cell where it can be replicated. A nucleic acid sequence can be "exogenous," which means that it is foreign to the cell into which the vector is being introduced or that the sequence is homologous to a sequence in the cell but in a position within the host cell nucleic acid in which the sequence is ordinarily not found. Vectors include plasmids, cosmids, viruses (bacteriophage, animal viruses, and plant viruses), and artificial chromosomes (*e.g.*, YACs). One of skill in the art would be well equipped to construct a vector through standard recombinant techniques (see, for example, Sambrook *et al.*, 2001 and Ausubel *et al.*, 1994, both incorporated herein by reference).

The term "expression vector" refers to any type of genetic construct comprising a nucleic acid coding for a RNA capable of being transcribed. Expression vectors can contain a variety of "control sequences," which refer to nucleic acid sequences necessary for the transcription and possibly translation of an operable linked coding sequence in a particular host cell. In addition to control sequences that govern transcription (promoters and enhancers) and translation, vectors and expression vectors may contain nucleic acid sequences that serve other functions as well that are well known to those of skill in the art, such as screenable and selectable markers, ribosome binding site, multiple cloning sites, splicing sites, poly A sequences, origins of replication, and other sequences that allow expression in different hosts.

Numerous expression systems exist that comprise at least a part or all of the compositions discussed above. Prokaryote- and/or eukaryote-based systems can be employed for use with the present invention to produce nucleic acid sequences, or their cognate polypeptides, proteins and peptides. Many such systems are commercially and widely available.

The nucleotide and protein, polypeptide and peptide sequences for various genes have been previously disclosed, and may be found at computerized databases known to those of ordinary skill in the art. For example, the nucleotide sequences of rRNAs of various organisms are readily available. One such database is the National Center for Biotechnology Information's Genbank and GenPept databases (http://www.ncbi.nlm.nih.gov/). The coding regions for all or part of these known genes may be amplified and/or expressed using the techniques disclosed herein or by any technique that would be know to those of ordinary skill in the art.

13. Nucleic Acid Arrays

Because the present invention provides efficient methods of enriching in mRNA, which can be used to make cDNA, the present invention extends to the use of cDNAs with arrays. The term "array" as used herein refers to a systematic arrangement of nucleic acid. For example, a cDNA population that is representative of a desired source (e.g., human adult brain) is divided up into the minimum number of pools in which a desired screening procedure can be utilized to detect a cDNA and which can be distributed into a single multi-well plate. Arrays may be of an aqueous suspension of a cDNA population obtainable from a desired mRNA source, comprising: a multi-well plate containing a plurality of individual wells, each individual well containing an aqueous suspension of a different content of a cDNA population. The cDNA population may include cDNA of a predetermined size. Furthermore, the cDNA population in all the wells of the plate may be representative of substantially all mRNAs of a predetermined size from a source. Examples of arrays, their uses, and implementation of them can be found in U.S. Patent Nos. 6,329,209, 6,329,140, 6,324,479, 6,322,971, 6,316,193, 6,309,823, 5,412,087, 5,445,934, and 5,744,305, which are herein incorporated by reference.

The number of cDNA clones array on a plate may vary. For example, a population of cDNA from a desired source can have about 200,000-6,000,000 cDNAs, about 200,000-2,000,000, 300,000-700,000, about 400,000-600,000, or about 500,000 cDNAs, and combinations thereof. Such a population can be distributed into a small set of multi-well plates, such as a single 96-well plate or a single 384-well plate. For instance, when about 1000-10,000 cDNAs, preferably about 3,500-7,000, more preferably about 5,000, from a population are present in a single well of a 96-well or 384-well plate, PCR can be utilized to clone a single, target gene using a set of primers.

The term a "nucleic acid array" refers to a plurality of target elements, each target element comprising one or more nucleic acid molecules immobilized on one or more solid surfaces to which sample nucleic acids can be hybridized. The nucleic acids of a target element can contain sequence(s) from specific genes or clones, e.g. from the regions identified here. Other target elements will contain, for instance, reference sequences. Target elements of various dimensions can be used in the arrays of the invention. Generally, smaller, target elements are preferred. Typically, a target element will be less than about 1 cm in diameter. Generally element sizes are from 1 μm to about 3 mm, between about 5 μm and about 1 mm. The target elements

of the arrays may be arranged on the solid surface at different densities. The target element densities will depend upon a number of factors, such as the nature of the label, the solid support, and the like. One of skill will recognize that each target element may comprise a mixture of nucleic acids of different lengths and sequences. Thus, for example, a target element may contain more than one copy of a cloned piece of DNA, and each copy may be broken into fragments of different lengths. The length and complexity of the nucleic acid fixed onto the target element is not critical to the invention. One of skill can adjust these factors to provide optimum hybridization and signal production for a given hybridization procedure, and to provide the required resolution among different genes or genomic locations. In various embodiments, target element sequences will have a complexity between about 1 kb and about 1 Mb, between about 10 kb to about 500 kb, between about 200 to about 500 kb, and from about 50 kb to about 150 kb.

Microarrays are known in the art and consist of a surface to which probes that correspond in sequence to gene products (e.g., cDNAs, mRNAs, cRNAs, polypeptides, and fragments thereof), can be specifically hybridized or bound at a known position. In one embodiment, the microarray is an array (i.e., a matrix) in which each position represents a discrete binding site for a product encoded by a gene (e.g., a protein or RNA), and in which binding sites are present for products of most or almost all of the genes in the organism's genome. In a preferred embodiment, the "binding site" (hereinafter, "site") is a nucleic acid or nucleic acid analogue to which a particular cognate cDNA can specifically hybridize. The nucleic acid or analogue of the binding site can be, e.g., a synthetic oligomer, a full-length cDNA, a less-than full length cDNA, or a gene fragment.

A microarray may contains binding sites for products of all or almost all genes in the target organism's genome, but such comprehensiveness is not necessarily required. Usually the microarray will have binding sites corresponding to at least about 50% of the genes in the genome, often at least about 75%, more often at least about 85%, even more often more than about 90%, and most often at least about 99%. Preferably, the microarray has binding sites for genes relevant to the action of a drug of interest or in a biological pathway of interest. A "gene" is identified as an open reading frame (ORF) of preferably at least 50, 75, or 99 amino acids from which a messenger RNA is transcribed in the organism (e.g., if a single cell) or in some cell in a multicellular organism. The number of genes in a genome can be estimated from the number

of mRNAs expressed by the organism, or by extrapolation from a well-characterized portion of the genome. When the genome of the organism of interest has been sequenced, the number of ORFs can be determined and mRNA coding regions identified by analysis of the DNA sequence.

The nucleic acid or analogue are attached to a solid support, which may be made from glass, plastic (*e.g.*, polypropylene, nylon), polyacrylamide, nitrocellulose, or other materials. A preferred method for attaching the nucleic acids to a surface is by printing on glass plates, as is described generally by Schena *et al.*, 1995a. *See also* DeRisi *et al.*, 1996; Shalon *et al.*, 1996; Schena *et al.*, 1995b. Each of these articles is incorporated by reference in its entirety.

Other methods for making microarrays, e.g., by masking (Maskos *et al.*, 1992), may also be used. In principal, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook et al., 1989, which is incorporated in its entirety for all purposes), could be used, although, as will be recognized by those of skill in the art, very small arrays will be preferred because hybridization volumes will be smaller.

Labeled cDNA is prepared from mRNA by oligo dT-primed or random-primed reverse transcription, both of which are well known in the art (see e.g., Klug *et al.*, 1987). Reverse transcription may be carried out in the presence of a dNTP conjugated to a detectable label, most preferably a fluorescently labeled dNTP. Alternatively, isolated mRNA can be converted to labeled antisense RNA synthesized by in vitro transcription of double-stranded cDNA in the presence of labeled dNTPs (Lockhart et al., 1996, which is incorporated by reference in its entirety for all purposes). In alternative embodiments, the cDNA or RNA probe can be synthesized in the absence of detectable label and may be labeled subsequently, e.g., by incorporating biotinylated dNTPs or rNTP, or some similar means (e.g., photo-cross-linking a psoralen derivative of biotin to RNAs), followed by addition of labeled streptavidin (e.g., phycoerythrin-conjugated streptavidin) or the equivalent.

Fluorescently-labeled probes can be used, including suitable fluorophores such as fluorescein, lissamine, phycoerythrin, rhodamine (Perkin Elmer Cetus), Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, FluorX (Amersham) and others (see, e.g., Kricka, 1992). It will be appreciated that pairs of fluorophores are chosen that have distinct emission spectra so that they can be easily distinguished. In another embodiment, a label other than a fluorescent label is used. For example, a radioactive label, or a pair of radioactive labels with distinct emission spectra, can be used (see

Zhao et al., 1995; Pietu et al., 1996). However, because of scattering of radioactive particles, and the consequent requirement for widely spaced binding sites, use of radioisotopes is a less-preferred embodiment.

In one embodiment, labeled cDNA is synthesized by incubating a mixture containing 0.5 mM dGTP, dATP and dCTP plus 0.1 mM dTTP plus fluorescent deoxyribonucleotides (e.g., 0.1 mM Rhodamine 110 UTP (Perken Elmer Cetus) or 0.1 mM Cy3 dUTP (Amersham)) with reverse transcriptase (*e.g.*, SuperScript™, Invitrogen Inc.) at 42°C for 60 min.

III. Methods for Isolating and Depleting Targeted Nucleic Acids

Methods of the invention involve preparing a sample comprising a targeted nucleic acid, preparing a bridging nucleic acid, preparing a capture nucleic acid, incubating the sample with the bridging nucleic acid, incubating the sample with a capture nucleic acid, incubating the bridging nucleic acid with the capture nucleic acid, incubating compounds under conditions allowing for hybridization among complementary regions, washing the sample and/or the capture and/or bridging nucleic acids, and isolating the capture nucleic acids and any accompanying compounds (compounds that bind or hybridize directly or indirectly to the capture nucleic acids). Steps of the invention are not required to be in a particular order and thus, the invention covers methods in which the order of the steps varies.

Hybridization conditions are discussed earlier. Wash conditions may involve temperatures between 20°C and 75°C, between 25°C and 70°C, between 30°C and 65°C, between 35°C and 60°C, between 40°C and 55°C, between 45°C and 50°C, or at temperatures within the ranges specified.

Buffer conditions for hybridization of nucleic acid compositions are well known to those of skill in the art. It is specifically contemplated that isostabilizing agents may be employed in hybridization and wash buffers in methods of the invention. U.S. Ser. No. 09/854,412 describes the use of tetramethylammonium chloride (TMAC) and tetraethylammonium chloride (TEAC) in such buffers; this application is specifically incorporated by reference herein. The concentration of an isostabilizing agent in a hybridization (binding) buffer may be between about 1.0 M and about 5.0 M, is about 4.0 M, or is about 2.0 M. Also specifically contemplated is a wash solution with an isostabilizing agent concentration of between about 0.1 M and 3.0 M, including 0.1 M increments within the range. Wash buffers may or may not contain Tris. However, in

some embodiments of the invention, the wash solution consists of water and no other salts or buffers. In some embodiments of the invention, the hybridizing or wash buffer may include guanidinium isothiocyanate, though in some embodiments this chemical is specifically contemplated to be absent. The concentration of guanidinium may be between about 0.4 M and about 3.0 M

A solution or buffer to elute targeted nucleic acids from the hybridizing nucleic acids (indirect or direct) may be implemented in some kits and methods of the invention. The elution buffer or solution can be an aqueous solution lacking salt, such as TE or water. Elution may occur at room temperature or it may occur at temperatures between 15°C and 100°C, between 20°C and 95°C, between 25°C and 90°C, between 30°C and 85°C, between 35°C and 80°C, between 40°C and 75°C, between 45°C and 70°C, between 50°C and 65°C, between 55°C and 60°C, or at temperatures within the ranges specified.

A. Quantitation of RNA

1. Assessing RNA yield by UV absorbance

The concentration and purity of RNA can be determined by diluting an aliquot of the preparation (usually a 1:50 to 1:100 dilution) in TE (10 mM Tris-HCl pH 8, 1 mM EDTA) or water, and reading the absorbance in a spectrophotometer at 260 nm and 280 nm.

An A_{260} of 1 is equivalent to 40 μg RNA/ml. The concentration (μg/ml) of RNA is therefore calculated by multiplying the A_{260} X dilution factor X 40 μg/ml. The following is a typical example:

The typical yield from 10 μg total RNA is 3 - 5 μg. If the sample is re-suspended in 25 μl, this means that the concentration will vary between 120 ng/μl and 200 ng/μl. One μl of the prep is diluted 1:50 into 49 μl of TE. The A_{260} = 0.1. RNA concentration = 0.1 X 50 X 40 μg/ml = 200 μg/ml or 0.2 μg/μl. Since there are 24 μl of the prep remaining after using 1 μl to measure the concentration, the total amount of remaining RNA is 24 μl X 0.2 μg/μl = 4.8 μg.

2. Assessing RNA yield with RiboGreen®

Molecular Probes' RiboGreen® fluorescence-based assay for RNA quantitation can be employed to measure RNA concentration.

B. Denaturing Agarose Gel Electrophoresis

Many mRNAs form extensive secondary structure. Ribosomal RNA depletion may be evaluated by agarose gel electrophoresis. Because of this, it is best to use a denaturing gel system to analyze RNA samples. A positive control should be included on the gel so that any unusual results can be attributed to a problem with the gel or a problem with the RNA under analysis. RNA molecular weight markers, an RNA sample known to be intact, or both, can be used for this purpose. It is also a good idea to include a sample of the starting RNA that was used in the enrichment procedure.

Ambion's NorthernMax™ reagents for Northern Blotting include everything needed for denaturing agarose gel electrophoresis. These products are optimized for ease of use, safety, and low background, and they include detailed instructions for use. An alternative to using the NorthernMax reagents is to use a procedure described in "Current Protocols in Molecular Biology", Section 4.9 (Ausubel et al., eds.), hereby incorporated by reference. It is more difficult and time-consuming than the Northern-Max method, but it gives similar results.

C. Agilent 2100 Bioanalyzer

1. Evaluating rRNA Removal with the RNA 6000 LabChip

An effective method for evaluating rRNA removal utilizes RNA analysis with the Caliper RNA 6000 LabChip Kit and the Agilent 2100 Bioanalyzer. Follow the instructions provided with the RNA 6000 LabChip Kit for RNA analysis. This system performs best with RNA solutions at concentrations between 50 and 250 ng/μl. Loading 1 μl of a typical enriched RNA sample is usually adequate for good performance.

2. Expected Results

In enriched mRNA samples from prokaryotes, the 16S and 23S rRNA peaks will be absent or present in only very small amounts. The peak calling feature of the software may fail to identify the peaks containing small quantities of leftover 16S and 23S rRNAs. A peak corresponding to 5S and tRNAs may be present depending on how the total RNA was initially purified. If RNA was purified by a glass fiber filter method prior to enrichment, this peak will be smaller. The size and shape of the 5S rRNA-tRNA peak is unchanged by some embodiments.

IV. KITS

Any of the compositions described herein may be comprised in a kit. In a non-limiting example, a bridging nucleic acid and a capture nucleic acid may be comprised in a kit. The kits will thus comprise, in suitable container means, a bridging nucleic acid and a capture nucleic of the present invention. It may also include one or more buffers, such as hybridization buffer or a wash buffer, compounds for preparing the sample, and components for isolating the capture nucleic acid via the nonreacting structure. Other kits of the invention may include components for making a nucleic acid array, and thus, may include, for example, a solid support.

The kits may comprise suitably aliquoted nucleic acid compositions of the present invention, whether labeled or unlabeled, as may be used to isolate, deplete, or separate a targeted nucleic acid. The components of the kits may be packaged either in aqueous media or in lyophilized form. The container means of the kits will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which a component may be placed, and preferably, suitably aliquoted. Where there are more than one component in the kit (bridging and capture nucleic acids may be packaged together), the kit also will generally contain a second, third or other additional container into which the additional components may be separately placed. However, various combinations of components may be comprised in a vial. The kits of the present invention also will typically include a means for containing the nucleic acids, and any other reagent containers in close confinement for commercial sale. Such containers may include injection or blow-molded plastic containers into which the desired vials are retained.

When the components of the kit are provided in one and/or more liquid solutions, the liquid solution is an aqueous solution, with a sterile aqueous solution being particularly preferred.

However, the components of the kit may be provided as dried powder(s). When reagents and/or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent. It is envisioned that the solvent may also be provided in another container means.

The container means will generally include at least one vial, test tube, flask, bottle, syringe and/or other container means, into which the nucleic acid formulations are placed,

preferably, suitably allocated. The kits may also comprise a second container means for containing a sterile, pharmaceutically acceptable buffer and/or other diluent.

The kits of the present invention will also typically include a means for containing the vials in close confinement for commercial sale, such as, *e.g.*, injection and/or blow-molded plastic containers into which the desired vials are retained.

Such kits may also include components that facilitate isolation of the targeting molecule, such as filters, beads, or a magnetic stand. Such kits generally will comprise, in suitable means, distinct containers for each individual reagent or solution as well as for the targeting agent.

A kit will also include instructions for employing the kit components as well the use of any other reagent not included in the kit. Instructions may include variations that can be implemented.

Kits of the invention may also include one or more of the following, in addition to a capture nucleic acid and a bridging nucleic acid:

1) Control RNA (*E. coli* or other appropriate RNA);
2) Nuclease-free water;
3) RNase-free containers, such as 1.5 ml tubes;
4) RNase-free elution tubes;
5) glycogen;
6) ethanol;
7) sodium acetate;
8) ammonium acetate;
9) magnetic stand or other magnetic field;
10) agarose;
11) nucleic acid size marker;
12) RNase-free tube tips;
13) and RNase or DNase inhibitors.

IV. Examples

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Furthermore, these examples are provided as one of many ways of implementing the claimed method and using the compositions of the invention. It is contemplated that the invention is not limited to the specific conditions set forth below, but that the conditions below provide examples of how to implement the invention.

EXAMPLE 1:

Materials

The following materials were used in the methods described herein for the selective removal of 16S and 23S rRNA and/or 18S and 28S rRNA, and hence mRNA enrichment, from total RNA. All steps are performed at room temperature unless otherwise indicated.

1. Bridging Nucleic Acids

In the following examples, the bridging regions are the poly-A stretches in the respective oligonucleotides.

Targeting regions for prokaryotic 16S and 23S rRNAs were designed based on a sequence comparison of different rRNAs from different bacteria to *E. coli* rRNA with MegAlign sequence analysis software version 4.05 from DNA Star, Incorporated (FIG. 2). The targeting regions are shown, in the examples below, 3' of the bridging regions. Thus, the targeting region encompasses the remaining, non-bridging region of each molecule described below. SEQ ID NOs are provided for the targeting regions of the bridging nucleic acids provided below (*i.e.*, sequence of bridging regions not included in SEQ ID NO.).

16S prokaryotic rRNA bridging oligonucleotides

d16S-358 (SEQ ID NO:1)

5'-AAAAAAAAAAAAAAAAAAACTGCTGCCTCCCGTAGGAGTCT-3'

d16S-537 (SEQ ID NO:2)
5'-AAAAAAAAAAAAAAAAAAACGTATTACCGCGGCTGCTGGCAC-3'

d16S-548 (SEQ ID NO:3)
5'-AAAAAAAAAAAAAAAAAAACGCCCAGTAATTCCGATTAACGC-3'

d16S-807 (SEQ ID NO:4)
5'-AAAAAAAAAAAAAAAAAATGGACTACCAGGGTATCTAATCC-3'

d16S-1092 (SEQ ID NO:5)
5'-AAAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTGCGGGACTT-3'

d16S-3' (SEQ ID NO:6)
5'-AAAAAAAAAAAAAAAAAAATAAGGAGGTGATCCAACCGCAGG-3'

23S prokaryotic rRNA bridging oligonucleotides

d23S-488 (SEQ ID NO:7)
5'-AAAAAAAAAAAAAAAAAAAGGTTCTTTTTCACTCCCCTCGCC-3'

d23S-581 (SEQ ID NO:8)
5'-AAAAAAAAAAAAAAAAAAAGACCCATTATACAAAAGGTACGC-3'

d23S-1118 (SEQ ID NO:9)
5'-AAAAAAAAAAAAAAAAAAAGCCCCGTTACATCTTCCGCGCAG-3'

d23S-1926 (SEQ ID NO:10)
5'-AAAAAAAAAAAAAAAAAAACGACAAGGAATTTCGCTACCTTA-3'

d23S-1954 (SEQ ID NO:11)
5'-AAAAAAAAAAAAAAAAAAACTTACCCGACAAGGAATTTCGC-3'

d23S-2511 (SEQ ID NO:12)
5'-AAAAAAAAAAAAAAAAAAAGAGCCGACATCGAGGTGCCAAAC-3'

d23S-3' (SEQ ID NO:13)
5'-AAAAAAAAAAAAAAAAAAAAGGTTAAGCCTCACGGTTCATT-3'

d23S-1704 (SEQ ID NO:15)
5'-AAAAAAAAAAAAAAAAAAACCCCTTCTCCCGAAGTTACGGGG-3'

d23S-1105 (SEQ ID NO:16)
5'-AAAAAAAAAAAAAAAAAAAGTGAGCTATTACGCTTTCTTT-3'

RNA oligo bridging oligonucleotide

r23S-3' (SEQ ID NO:14)
5'-AAAAAAAAAAAAAAAAAAAAAGGUUAAGCGUCACGGUUCAUU-(inverted (dT))-3'
(inverted refers to bases attached 3' to 3')

Eukaryotic 18S rRNA bridging oligonucleotides

d18S-3711 (SEQ ID NO:17)
AAA AAA AAA AAA AAA AAA TAC CGG CCG TGC GTA CTT AGA CA

d18S-4238 (SEQ ID NO:18)
AAA AAA AAA AAA AAA AAA TGC CCT CCA ATG GAT CCT CGT TA

d18S-5482 (SEQ ID NO:19)
AAA AAA AAA AAA AAA AAA CTA CGG AAA CCT TGT TAC GAC TT

Eukaryotic 28S rRNA bridging oligonucleotides

d28S-11599 (SEQ ID NO:20)
AAA AAA AAA AAA AAA AAA GAG CAC TGG GCA GAA ATC ACA TC

d28S-7979 (SEQ ID NO:21)
AAA AAA AAA AAA AAA AAA GTT TCT TTT CCT CCG CTG ACT AA

d28S-12533 (SEQ ID NO:22)
AAA AAA AAA AAA AAA AAA TCC TCA GCC AAG CAC ATA CAC CA

2. *Binding Buffer* (also referred to as hybridization buffer)
3 M TMAC, 10 mM Tris, (pH 7.0)

3. *Bridging Nucleic Acid Mixture*
Mixtures of 16S, 23S, 18S, and/or 28S bridging oligonucleotides were used. All oligonucleotides were purchased from IDT and purified from polyacrylamide gels.

4. *Capture Nucleic Acid (Oligo(dT) MagBeads)*
Seradyn MGOL #2815-2103.

5. *Wash Solution*
2 M TMAC, 6.67 mM Tris (pH 7.0) (this is a dilution of binding buffer).

EXAMPLE 2:
Methods for rRNA Depletion from Prokaryotic Total RNA

The following methods are provided by way of example for practicing methods of the invention. They have been performed and shown to effect methods of the invention. The invention is not intended to be limited to these protocols, and it is specifically contemplated that variations of the methods below may be employed that fall within the scope of the invention if they effect depletion, isolation, or separation of a targeted nucleic acid, particularly rRNA.

This example demonstrates the depletion of 16S and 23S rRNA from *E. coli* total RNA.

RNA/Bridging Nucleic Acid Mixture Annealing

RNA (10 µg/15 µl) was added to 200 µl of binding buffer. The bridging nucleic acid mixture consisted of d16S-807 (5 µM), d16S-1092 (5 µM), d23S-1954 (5 µM), d23S-2511 (5 µM). The bridging nucleic acid mixture (4 µl) was added to the RNA and the mixture was incubated at 70°C for 10 minutes and then shifted to 37°C for 30 minutes.

Thirty minutes was found to be an adequate time for the annealing step. Longer time periods can be used with no adverse effects. Between fifteen and 120 minutes have been used successfully in the methods of the invention.

Preparation of Capture Nucleic Acid

Capture nucleic acid (Oligo (dT) MagBeads, Seradyn) in storage buffer was mixed and 50 µl was removed to a separate tube. A magnetic stand was applied to the side of the tube to capture the magnetic beads and the supernatant was removed. The capture nucleic acid was equilibrated one time with distilled, deionized water (50 µl) and once with binding buffer (50 µl). The captured nucleic acid was captured again with a magnetic stand, and the binding buffer wash was removed. The magnetic beads were resuspended in 50 µl of binding buffer.

rRNA Capture

Following the 30 minute annealing of RNA with the bridging nucleic acid mixture, the capture nucleic acid was added and the mixture was incubated at room temperature for 15 minutes. A magnetic stand was then applied to the tube to capture the magnetic beads. The supernatant containing mRNA, 5S rRNA, and tRNAs was removed to another tube and saved. An optional washing step was performed next. The magnetic beads were washed with Wash

Solution (100 µl) and captured again. The wash supernatant was removed and added to the original supernatant.

Fifteen minutes was found to be an adequate time for rRNA capture. Longer time periods can be used with no adverse effects. rRNA capture likely occurs rapidly, and capture times of 5 minutes – 60 minutes have been used successfully in the methods of the invention.

Precipitating mRNA

mRNA, 5S rRNA, and tRNAs were precipitated by adding 1/10 volume of 3M NaOAc (pH 5.5) and 3 volumes of 100% EtOH and incubating at -20^{o}C for 60 minutes. The precipitated RNA was pelleted in a microfuge, washed with 70% EtOH, and resuspended in TE (pH 8.0).

Analysis of Purified mRNA

Purified mRNA was analyzed with the Caliper RNA 6000 LabChip kit on an Agilent Bioanalyzer. Purified RNA was compared with a control *E. coli* total RNA sample that was carried through the reaction as described above, except that the Bridging Nucleic Acid Mixture was left out. This assay system uses electrophoretic and electrokinetic separation in a capillary electrophoresis type system. The rRNAs appear as peaks on an electropherogram (FIG. 3). The percentage of a rRNA present in the sample is calculated from the area under the peak.

Under the protocol conditions described above, the 5S + tRNA peak area is essentially the same in the control and in experimental samples. The % of 16S or 23S rRNA removed was calculated using the ratios of $16S_{peak\ area}/5S_{peak\ area}$ and $23S_{peak\ area}/5S_{peak\ area}$. Enriched and control RNAs with similar 5S + tRNA peak areas were compared.

$$\%\ 16S\ rRNA\ removed = \frac{(16S_{peak\ area}/5S_{peak\ area})_{no\ oligos\ control} - (16S_{peak\ area}/5S_{peak\ area})_{experimental}}{(16S_{peak\ area}/5S_{peak}\ area)_{no\ oligos\ control}} X\ 100$$

A corresponding formula was used to calculate % 23S rRNA removed.

Electropherograms of RNA from a control reaction and from an experimental reaction after ribosomal RNA depletion are shown in FIG. 3 and FIG. 4.

EXAMPLE 3:

Evaluations of Efficacy with Prokaryotic Targets

The materials and methods of Examples 1 and 2 were employed to determine the efficiency of removal of 16S rRNA or 23S rRNA or both from *E. coli* total RNA. Changes in the parameters of the experiments are noted when appropriate. These experiments were performed to evaluate the efficacy of various bridging nucleic acids and reaction conditions.

The following results are from reactions that employed 10 µg of *E. coli* total RNA, 40 pmol of total 16S rRNA bridging nucleic acid, 40 pmol of total 23S rRNA bridging nucleic acid, and 50 µl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-358/d23S-2511	96.48285	89.86496
d16S-537/d23S-1954	97.47974	91.32074
d16S-537/d23S-2511	97.48704	91.216
d16S-807/d23S-1954	95.79126	89.85388
d16S-807/d23S-2511	95.25362	91.06399
d16S-1092/d23S-1118	97.91265	96.50658
d16S-1092/d23S-1954	96.7473	89.40605
d16S-1092/d23S-2511	97.61689	91.5964
d16S-358/d23S-1954	96.74434	88.07242
d16S1092/d23S-1954 (20 pmol) d23S-2511 (20 pmol)	97.19134	98.44728

The following results are from reactions that employed 10 µg of *E. coli* total RNA, 26 pmol of 16S rRNA bridging nucleic acid, 26 pmol of 23S rRNA bridging nucleic acid, and 35 µl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-1092/d23S-1118	97.38534	95.02083
d16S-1092/d23S-1957	97.8291	90.798

The following results are from reactions that employed 10 μg of *E. coli* total RNA, 75 pmol of 16S rRNA bridging nucleic acid, 75 pmol of 23S rRNA bridging nucleic acid, and 100 μl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S.....23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-1092.....d23S-1118	99.14812	99.11895
d16S-1092.....d23S-1954	98.79938	98.45245
d16S-1092.....d23S-2511	99.00567	98.84033

The following results are from reactions that employed 10 μg of *E. coli* total RNA, 37.5 pmol of 16S rRNA bridging nucleic acid, 37.5 pmol of 23S rRNA bridging nucleic acid, and 50 μl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-1092/d23S-1118	98.95563	98.28748
d16S-1092/d23S-1954	97.83593	94.84438

The following results are from reactions that employed 10 μg of *E. coli* total RNA, 75 pmol of 16S rRNA bridging nucleic acid or 75 pmol of 23S rRNA bridging nucleic acid with 75 μl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	% 16S Removed	% 23S Removed
n.a./d23S-581	-	98.98529
n.a./d23S-581	-	98.87251
n.a./d23S-1118	-	93.62175
n.a./d23S-1118	-	91.4927
n.a./d23S-1954	-	98.68262
n.a./d23S-1954	-	99.03237
n.a./d23S-2511	-	99.31982
n.a./d23S-2511	-	99.13291
d16S-358/n.a.	97.65586	-
d16S-358/n.a.	97.51393	-
d16S-537/n.a.	99.16427	-
d16S-537/n.a.	98.92345	-
d16S-807/n.a.	98.0661	-
d16S-807/n.a.	98.14292	-

n.a. = not applicable

The following results are from reactions that employed 5 μg of *E. coli* total RNA, 25 pmol of each 16S rRNA or 23S rRNA bridging nucleic acid, and 25 μl of capture nucleic acid described in Example 1. The rRNA/bridging nucleic acid annealing reaction was for 60 minutes at 37°C.

Bridging Nucleic Acid 16S/23S	% 16S Removed	% 23S Removed
n.a./d23S-488	-	~100
n.a./d23S-1118	-	~100
d16S-3'/d23S-488	89.024	94.228
d16S-548/d23S-488	~100	93.718
d16S-1092/d23S-488	~100	92.652

The following results are from reactions that employed 5 μg of *E. coli* total RNA, 16S rRNA bridging nucleic acid as indicated, 23S rRNA bridging nucleic acid as indicated, and 25 μl of capture nucleic acid described in Example 1. The rRNA/bridging nucleic acid annealing reaction was for 120 minutes at 37°C.

Bridging Nucleic Acid 16S/23S	% 16S Removed	% 23S Removed
d16S-3' (25 pmol)/n.a.	89.137	-
d16S-548 (25 pmol)/n.a.	~100	-
d16S-1092 (25 pmol)/n.a.	~100	-
d16S-3' (25 pmol) d16S-548 (25 pmol)/n.a.	~100	-
d16S-3' (25 pmol) d16S-1092 (25 pmol)/n.a.	~100	-
d16S-548 (25 pmol) d16S-1092 (25 pmol)n.a.	~100	-
d16S-548 (25 pmol)/ d23S-3' (25 pmol)	~100	~100
d16S-1092 (25 pmol)/ d23S-3' (25 pmol)	~100	~100
d16S-3' (25 pmol)/ d23S-3' (25 pmol)	92	~100

EXAMPLE 4:
The Effect of Washing the Capture Nucleic Acid

The purpose of this experiment was to determine if washing the capture nucleic acid and combining the wash with the purified mRNA had an effect on the presence of rRNA in the purified mRNA sample. Reactions employed 10 μg of *E. coli* total RNA, 75 pmol d16S-1092, 75 pmol of d23S-d1118, and 100 μl of capture nucleic acid described in Example 1. The rRNA/bridging nucleic acid annealing reaction proceeded for 60 min at 37°C. After the nucleic acid capture step, the capture nucleic acid (with bound rRNA) was resuspended and washed with 100 μl of the indicated solution at room temperature for 5 minutes. The capture nucleic acid was re-captured with a magnetic stand and the supernatant was removed and combined with mRNA in the supernatant from the first capture. mRNA in the combined supernatants were precipitated with ethanol and evaluated with RNA 6000 Lab Chip assay for the presence of rRNAs. The percent of rRNA removal for the entire process is indicated in the table below.

Wash	% 16S Removed	% 23S Removed
0.4 M TMAC	66.061	66.175
1.0 M TMAC	95.810	96.708
1.5 M TMAC	~100	~100
2.0 M TMAC	~100	~100

These results demonstrate that lowering the molarity of the TMAC wash solution increases the stringency of the rRNA capture reaction when the temperature is held constant at room temperature. The results also demonstrate that washing the capture nucleic acid magnetic beads with 1.5 and 2.0 M TMAC does not remove rRNA from the capture nucleic acid.

EXAMPLE 5:
Evaluation of Efficacy with Prokaryotic and Eukaryotic rRNA Targets

The purpose of this example was to evaluate efficacy of the methods of the invention for depleting 16S rRNA, 18S rRNA, 23S rRNA, and 28S rRNA from mixtures of prokaryotic and eukaryotic total RNA. Depletion methods were verified using various mammalian samples, including rat livers.

Equal amounts (2.5 μg) of *E. coli* total RNA and rat liver total RNA were mixed prior to the mRNA enrichment procedure. The bridging oligonucleotides employed were:

d16S-1092 (10 pmol)
d16S- 807 (10 pmol)
d23S-1954 (10 pmol)
d23S-2511 (10 pmol)
d18S-3711 (20 pmol)
d28S-11599 (20 pmol)

The reaction used 50 µl of capture nucleic acid as described in Example 1. No wash step was employed. Otherwise the reaction was performed according to methods in Example 2. The results are shown in FIG. 5A and 5B. Note that all rRNAs were depleted except the 5S and 5.8S rRNAs for which no bridging oligonucleotides were added.

EXAMPLE 6:

Evaluation of Efficacy with Human rRNA Targets

Additional experiments were done using human samples to evaluate the extent of human rRNA depletion using the bridging oligonucleotides shown below. Depletion of 18S rRNA and 28S rRNA was observed from human liver total RNA. rRNAs were depleted from human liver total RNA (5 µg). The bridging oligonucleotides employed were:

d18S-3711 (40 pmol)
d28S-11599 (40 pmol)

The reaction used 50 µl of capture nucleic acid as described in Example 1. No wash step was employed. Otherwise the reaction was performed according to Example 2.

The results are shown in FIG. 6A and 6B. Note that all rRNAs (18S, 28S) were depleted except the 5S and 5.8S rRNAs for which no bridging oligonucleotides were added.

EXAMPLE 7:

Evaluation of Efficacy with Rat rRNA Targets

Additional experiments were done using rat samples to evaluate the extent of rat rRNA depletion using the bridging oligonucleotides shown below. Depletion of 18S rRNA and 28S rRNA was observed from rat liver total RNA. rRNAs were depleted from rat liver total RNA (5 µg). The bridging oligonucleotides employed were:

d18S-3711R-polyA (40 pmol)
d28S-11599R-polyA (40 pmol)

The reaction used 50 µl of capture nucleic acid as described in Example 1. No wash step was employed. Otherwise the reaction was performed according to Example 2.

The results are shown in FIG. 7A and 7B. Note that all rRNAs (18S, 28S) were depleted except the 5S and 5.8S rRNAs for which no bridging oligonucleotides were added.

EXAMPLE 8:

Evaluation of Efficacy with Mouse rRNA Targets

Additional experiments were done using mouse samples to evaluate the extent of rat rRNA depletion using the bridging oligonucleotides shown below. Depletion of 18S rRNA and 28S rRNA was observed from mouse liver total RNA (5 µg). The bridging oligonucleotides employed were:

d18S-3711R-polyA (40 pmol)
d28S-11599R-polyA (40 pmol)

The reaction used 50 µl of capture nucleic acid as described in Example 1. No wash step was employed. Otherwise the reaction was performed according to Example 2.

The results are shown in FIG. 8A and 8B. Note that all rRNAs (18S, 28S) were depleted except the 5S and 5.8S rRNAs for which no bridging oligonucleotides were added.

EXAMPLE 9:

Use of Purified *E. coli* mRNA in Gene Array Expression Analysis

mRNA was purified from total *E. coli* RNA (10 µg) using the methods of the invention as described in Example 2. A control reaction was also performed in which the bridging nucleic acid mixture was omitted form the reaction. Control total RNA and purified mRNA (1.5 µg) were added to 70 pmol random hexamers in a final volume of 7.25 µl. The mixture was heated at 70°C for 10 minutes, then transferred to ice for 3 minutes. The following components were added to each reaction:

5 µl	cDNA 1st strand synthesis buffer (Invitrogen)
2.5 µl	0.1 M DTT
1.25 µl	10 mM dATP

1.25 µl	10 mM dGTP
1.25 µl	10 mM dTTP
5 µl	10 mCi/ml ^{33}P-dCTP (Perkin Elmer-NEN)
1 µl	Superscript II reverse transcriptase (Invitrogen) 200 U/µl

The reactions were incubated at 42°C for 120 minutes. Unincorporated nucleotides were removed from the reactions with a Qiaquick PCR cleanup column (Qiagen). The labeled cDNAs (3×10^7 cpm/blot) were used to probe replicate portions of Panorama™ *E. coli* gene arrays, using hybridization buffers supplied by the array manufacturer (Sigma-Genosys). The arrays were washed and exposed to film. This example demonstrates a dramatic increase in hybridization signal (sensitivity) on gene arrays when labeled cDNA is prepared from bacterial mRNA, purified according to the methods of the invention, rather than from total RNA.

EXAMPLE 10:

Instructions for Use with Kit

The following instructions have been followed with a kit of the invention described below for the successful depletion of 16S and 23S rRNA from a sample comprising prokaryotic RNA populations. Bridging oligonucleotides with targeting regions complementary to 18S and 28S rRNA may be employed according to the method below to effect a similar result (as in Examples 5-8).

Materials Provided with a Kit Embodiment

30 µl	Control RNA
1.2 ml	Capture Nucleic Acid **[as in Example 1]**
7 ml	Binding Buffer **[as in Example 1]**
95 µl	Bridging Oligonucleotide Mix **[as in Example 2]**
2.4 ml	Wash Solution **[as in Example 1]**
1.75 ml	Nuclease-free Water
50 ea	RNase-free 1.5 ml tubes
25 ea	RNase-free 2ml Elution tubes
200 µl	Glycogen (5 mg/ml)
875 µl	3 M NaOAc

Experimental Parameters

A. RNA Source

This mRNA enrichment procedure is designed to work with purified total RNA from many different bacteria, including both gram-positive and gram-negative species. The procedure was optimized with total *E. coli* RNA and has been found to remove 90-99% of the rRNA from *Bacillus subtilis*, *Staphylococcus aureus*, *Prochlorococcus* sp., *Neisseria meningitidis*, and *Pseudomonas aeruginosa*, for example. It is contemplated that any eubacterial species may be targeted using the methods and compositions of the invention.

This procedure is designed so that small RNAs (including tRNA and 5S rRNA) remain in the enriched mRNA population. However, if the loss of very small RNA species (<200 base) will not be an issue, the initial isolation of total RNA should be performed with Ambion's RNAQUEOUS KIT. The RNAQUEOUS KIT will remove most small RNA species and provide the highest possible level of mRNA enrichment. If small RNAs are of interest to the user, it is best to avoid glass fiber filter-based purification.

B. Precipitate RNA to remove salt and concentrate if necessary

Total RNA prepared from a solid-phase extraction method such as RNAQUEOUS can be used immediately after elution because such samples are unlikely to have high levels of salt. On the other hand, RNA isolated by methods that include organic extractions, for example using the products RNAWIZ, TRIZOL or ToTALLY RNA, may have a substantial amount of residual salt. If RNA from these types of procedures has been precipitated only a single time, we recommend doing a second alcohol precipitation and 70% EtOH wash to remove residual salt before starting the enrichment procedure.

The recommended maximum amount of RNA per reaction is 10 µg and the recommended maximum volume for the RNA is 15 µl. If the RNA sample is too dilute, it will be necessary to precipitate and concentrate the RNA to at least 10 µg/15 µl. Precipitate the RNA with:

- 0.1 volume 5 M Ammonium Acetate or 3 M sodium acetate
- 1 µl Glycogen (The glycogen acts as a carrier to increase precipitation efficiency from dilute RNA solutions; it is unnecessary for solutions with 200 µg RNA/ml)
- 2.5 volumes 100% ethanol

a. Leave the precipitation mixture at –20°C overnight, or quick-freeze it in either ethanol and dry ice, or in a –70°C freezer for 30 minutes.

b. Recover the RNA by centrifugation at 12,000 x g for 30 minutes at 4°C.

c. Carefully remove and discard the supernatant. The RNA pellet may not adhere tightly to the walls of the tubes, so we suggest removing the supernatant by gentle aspiration with a fine-tipped pipette.

d. Centrifuge the tube briefly a second time, and aspirate any additional fluid that collects with a fine-tipped pipette.

e. Add 1 ml 70% ethanol, and vortex the tube a few times. Repellet the RNA by microcentrifuging, for 10 minutes at 4°C. Remove supernatant carefully as in steps c and d above.

RNA should be dissolved in TE or Ambion's THE RNA STORAGE SOLUTION. It is important to accurately quantitate RNA so as not to overload the system. Ambion recommends using the RiboGreen RNA Quantitation Assay and Kit (Molecular Probes) or a high quality, calibrated spectrophotometer.

C. Save an aliquot of your total RNA

If possible, retain a small aliquot (~1–2 µg) of the total RNA used for comparison with enriched mRNA by gel electrophoresis after the procedure is finished.

Instructions

A. Anneal RNA and Bridging Oligonucleotide Mix

1. Add RNA to Binding buffer

Add total RNA (up to 10 µg total RNA in a maximum volume of 15 µl) to 200 µl Binding Buffer in a 1.5 ml tube provided with the kit. Close the tube and tap or vortex gently to mix.

2. Add Bridging Oligonucleotide Mix to RNA

Add 4.0 µl of the Bridging Oligonucleotide Mix to the RNA in Binding Buffer. Close the tube and tap or vortex gently to mix. Pulse in a microcentrifuge very briefly to get mixture to bottom of tube.

3. Incubate reactions at 70°C for 10 minutes.

Incubating the mixture at 70°C for 10 minutes denatures secondary structures in RNA, including the 16S and 23S rRNAs, allowing for maximal hybridization of the bridging oligonucleotides to the rRNAs.

4. Incubate reactions at 37°C for 1 hour.

Incubating the mixture at 37°C for 1 hour allows for binding of the bridging oligonucleotides to the 16S and 23S rRNA. The Binding Buffer has been optimized to function specifically and efficiently at this temperature.

B. Prepare the Capture Nucleic Acid

During the 1 hour $RNA^{Bridging}$ Oligonucleotide Mix annealing step, prepare the Capture Nucleic Acid. The Capture Nucleic Acid is in a 1% (10 mg/ml) suspension, vortex the tube briefly before pipetting to be sure they are well suspended.

1. Aliquot the Capture Nucleic Acid

For each 10 μg reaction remove 50 μl Capture Oligos to a 1.5 ml tube. Capture Nucleic Acid for up to 10 reactions can be processed in a single 1.5 ml tube.

2. Wash the Capture Nucleic Acid once with water and once with Binding Buffer

a. Capture the beads (Capture Nucleic Acid) by placing the tube on the Magnetic Stand. Leave the tube on the stand until all of the Capture Nucleic Acid is arranged inside the tube near the magnet. This will take ~3 minutes for microfuge tubes.

b. Carefully remove the supernatant by aspiration, leaving the beads in the tube, and discard the supernatant.

c. Add Nuclease Free Water to the captured beads at a ratio of 50 μl/50 μl beads).

d. Remove the tube from the Magnetic Stand, resuspend the beads by gently vortexing briefly, recapture the beads with a Magnetic Stand, carefully aspirate the supernatant, leaving the beads in the tube, and discard the supernatant.

e. Add Binding Buffer to the captured beads at a ratio of 50 μl/50 μl beads).

f. Repeat step d.

3. Resuspend the Capture Nucleic Acid in Binding Buffer

a. Add Binding Buffer to the captured beads at a ratio of 50 µl/50 µl beads).

b. Remove the tube from the Magnetic Stand, resuspend the beads by gently tapping the tube or very gentle vortexing.

c. Pulse spin in a microcentrifuge to get liquid to the bottom of the tube.

C. Capture the rRNA with Capture Nucleic Acid and Recover the Enriched mRNA

1. Add Capture Nucleic Acid (50 µl/rxn) to RNA/Bridging Oligonucleotide Mix and incubate at RT for 15 minutes.

a After the 1 hour incubation at 37°C (Step A.4) remove tubes to room temperature (RT) and immediately add 50 µl of the washed and equilibrated beads (Capture Nucleic Acid, from Step B.3c) to each purification reaction . Very gently vortex or tap tube to mix briefly and pulse spin in a microcentrifuge to get liquid to the bottom of the tube.

b. Incubate 15 minutes at RT. During this step the oligonucleotide sequence on the Capture Nucleic Acid anneals to the bridging oligonucleotides. The bridging oligonucleotides remain hybridized to the 16S and 23S rRNAs. The hybridization "sandwich" of bridging oligonucleotide and capture oligonucleotide (via the capture region on the capture oligo and the bridging region on the bridging oligo) is formed at this step.

2. Recover the supernatant containing the enriched mRNA.

a. Capture the beads by placing the tube on the Magnetic Stand. Leave the tube on the stand until all of the beads are arranged inside the tube near the magnet. This will take ~3 minutes for microfuge tubes. Allow the beads to be completely captured by the magnet for at least 3 minutes.

b. Remove the supernatant by aspiration, being careful not to dislodge the beads. Put the supernatant into a 2 ml nipple bottom tube on ice and save. Do not be overly concerned if there seems to be beads in the removed supernatant. The excess can be removed at the end of the procedure. The supernatant contains the enriched mRNA sample.

3. Wash the Oligo MagBeads with Wash Solution and recover the wash.

a. Add Wash Solution to the captured beads at a ratio of 100 μl Wash Solution/50 μl beads.

b. Remove the tube from the Magnetic Stand, resuspend the beads by gently vortexing briefly.

c. Incubate at RT for 5 minutes.

d. Recapture the beads with the Magnetic Stand as in step C.2a. Allow the beads to be completely captured by the magnet for at least 3 minutes.

e. Remove the supernatant by aspiration, being careful not to dislodge the beads. Put this supernatant in the 2 ml nipple bottom tube on ice with that from step C.2b.

D. Precipitate and resuspend the enriched mRNA in the supernatant.

1. Perform an EtOH precipitation on the collected supernatant.

a. Add 1/10 Volume 3M NaOAc (35 μl) and 5 mg/ml glycogen to a final concentration of 100μg/ml (7 μl) to the supernatant from step C.3.e. (the supernatant volume should be ~350 μl).

b. Briefly vortex the sample to mix.

c. Add 3 Vol. ice cold 100% EtOH (1175 μl) and mix well by vortexing the sample.

d. Precipitate the sample at –20°C for at least 1 hour.

e. Centrifuge the sample for 30 min. @ 13,000 rpm.

f. Carefully decant the supernatant.

g. Add 750ml ice cold 70% EtOH, vortex briefly, and centrifuge for 5 min. @ 13,000 rpm. Decant the supernatant.

h. Repeat step D.1.g.

i. After decanting the supernatant spin briefly to collect. Remove the remaining supernatant with a pipettor, being careful not to dislodge the pellet. Air dry for 5 min.

2. Resuspend the enriched mRNA in an appropriate buffer.

a. After the pellet has air dried for no more than 5 min. add 2 μl TE pH 8.0 (RNA STORAGE SOLUTION, 1 mM EDTA or Nuclease-Free ddH_2O could be substituted).

b. Allow the RNA to resuspend for 15 min. at room temperature. Vortex the sample vigorously to resuspend. Collect the sample by brief centrifugation. NOTE: If the pellet refuses to go into solution the sample can be incubated for 5 min. @ 70°C. This should help resuspend the pellet. NOTE: Often there will be beads remaining in the sample after the precipitation (This will cause the RNA solution to appear brownish in color). This can be remedied by applying the sample to the Magnetic stand for ~3 min. and removing the supernatant to a new tube.

E. Compatibility with respect to other microorganisms

Based on experimental evidence and sequence information, the following organisms should be compatible (removal of 16S rRNA and of 23S rRNA) with the oligos identified in Example 1 (non-control oligos): *Acidithiobacillus ferrooxidans, Acinetobacter calcoaceticus, Actinobacillus actinomycetemcomitans, Aeromonas hydrophila, Agrobacterium tumefaciens, Alcaligenes faecalis, Bacillus alcalophilus, Bacillus anthracis, Bacillus cereus, Bacillus halodurans, Bacillus licheniformis, Bacillus mycoides, Bacillus subtilis, Bacillus thuringiensis, Bartonella bacilliformis, Bordetella avium, Bordetella bronchiseptica, Bordetella parapertussis, Bordetella pertussis, Borrelia burgdorferi, Bradyrhizobium japonicum, Bradyrhizobium lupini, Brevundimonas diminunata, Brucella melitensis, Brucella melitensis biovar suis, Buchnera aphidicola, Buchnera sp. APS, Burkholderia cepacia, Burkholderia mallei, Burkholderia pseudomallei, Caulobacter crescentus, Chlamydia muridarum, Chlamydia suis, Chlamydia trachomatis, Chlamydophila abortus, Chlamydophila caviae, Chlamydophila felis, Chlamydophila pecorum, Chlamydophila pneumoniae, Chlamydophila psittaci, Chlorobium limicola, Chlorobium tepidum, Citrobacter freundii, Clostridium acetobutylicum, Clostridium difficile, Clostridium histolyticum, Corynebacterium diptheriae, Corynebacterium glutamicum, Cytophaga hutchisonii, Desulfovibrio vulgaris, Dichelobacter nodosus, Enterococcus asini, Enterococcus avium, Enterococcus casseliflavus, Enterococcus cecorum, Enterococcus*

columbae, Enterococcus dispar, Enterococcus durans, Enterococcus faecalis, Enterococcus faecium, Enterococcus flavescens, Enterococcus gallinarum, Enterococcus hirae, Enterococcus malodoratus, Enterococcus mundtii, Enterococcus pseudoavium, Enterococcus raffinosus, Enterococcus saccharolyticus, Enterococcus sulfureus, Erwinia chrysanthemi, Escherichia coli, Fibrobacter succinogenes, Frankia sp., Fusobacterium nucleatum, Geobacillus stearothermophilus, Geobacter sulfurreducens, Gluconacetobacter europaeus, Gluconacetobacter intermedius, Gluconacetobacter xylinus, Haemophilus ducreyi, Haemophilus influenzae, Klebsiella pneumoniae, Lactobacillus amylolyticus, Lactobacillus delbrueckii, Lactococcus lactis, Leuconostoc carnosum, Leuconostoc lactis, Leuconostoc mesenteroides, Listeria grayi, Listeria innocua, Listeria ivanovii, Listeria monocytogenes, Listeria seeligeri, Melissococcus plutonius, Micrococcus luteus, Mycobacterium avium, Mycobacterium avium supsp. Paratuberculosis, Mycobacterium bovis, Mycobacterium kansasii, Mycobacterium leprae, Mycobacterium phlei, Mycobacterium smegmatis, Mycobacterium tuberculosis, Myxococcus xanthus, Neisseria gonorrhoeae, Neisseria meningitidis, Nitrosomonas europaea, Pasteurella multocida, Peptococcus niger, Plesiomonas shigelloides, Pseudomonas aeruginosa, Pseudomonas putida, Pseudomonas syringae, Ralstonia pickettii, Ralstonia solanacearum, Renibacterium salmoninarum, Rhizobium vitis, Rhodococcus erythropolis, Rhodococcus fascians, Rhodopseudomonas palustris, Rhodospirillum rubrum, Rickettsia akari, Rickettsia australis, Rickettsia bellii, Rickettsia canadensis, Rickettsia conorii, Rickettsia montanensis, Rickettsia parkeri, Rickettsia prowazekii, Rickettsia rhipicephali, Rickettsia rickettsii, Rickettsia sibirica, Rickettsia typhi, Salmonella bongori, Salmonella enterica, Salmonella enteritidis, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Shewanella putrefaciens, Sinorhizobium meliloti, Sporosarcina globispora, Staphylococcus aureus, Staphylococcus carnosus, Staphylococcus condimenti, Staphylococcus epidermidis, Stigmatella aurantiaca, Streptococcus equii, Streptococcus gordonii, Streptococcus macedonicus, Streptococcus mitis, Streptococcus mutans, Streptococcus oralis, Streptococcus parauberis, Streptococcus pneumoniae, Streptococcus pyogenes, Streptococcus thermophilus, Streptococcus uberis, Streptomyces ambofaciens, Streptomyces coelicolor, Streptomyces griseus, Streptomyces lividans, Streptomyces nodosus, Streptomyces rimosus, Thermoanaerobacter tengcongensis, Thermobifida fusca, Thermomonospora chromogena, Thiobacillus ferrooxidans, Trteponema denticola, Treponema pallidum, Vibrio cholerae, Yersinia enterocolitica, Yersinia pestis, Xanthomonas campestris, Xanthomonas axonopodis pv. Citri, and *Xylella fastidiosa.*

Based on experimental evidence and sequence information, the following organisms should be partially compatible (removal of 23S rRNA and of 50-100% 16S rRNA) with oligos identified in Example 1 (non-control): *Azotobacter vinelandii, Bacteroides fragilis, Carboxydothermus hydrogenoformans, Clostridium tyrobutyricum, Desulfovibrio vulgaris, Dictyoglomus thermophilum, Enterococcus solitarius, Erysipelothrix rhusiopathiae, Erysipelothrix tonsillarum, Flexibacter flexilis, Legionella pneumophila, Leptospira interrogans, Leucothrix mucor, Listeria welshimeri, Methylococcus capsulatus, Myroides odoratus, Oenococcus oeni, Paracoccus denitrificans, Pectinatus frisingensis, Porphyromonas gingivalis, Prevotella intermedia, Silicibacter pomeroyi, Tannerella forsythensis, Tetragenococcus halophilus, Thermobispora bispora, Thermus thermophilus* and, *Thiomonas cuprina*

Based on experimental evidence and sequence information, the following organisms should be partially compatible (removal of 16S rRNA and 50-100% of 23S rRNA) with oligos identified in Example 1 (non-control): *Burkholderia fungorum, Clostridium perfringens, Desulfitobacterium hafniense, Magnetospirillum magnetotacticum, Mesorhizobium loti, Nannocystis exedens, Novosphingobium aromaticivorans, Parachlamydia acanthamoebae, Ruminobacter amylophilus, Ruminococcus albus, Tropheryma whipplei,* and *Wolbachia endosymbiont* of *Drosophila melanogaster.*

Based on experimental data and sequence information, the following organisms are believed to be incompatible with the oligos in Example 1: *Archaebacteria, Campylobacter spp., Chloroflexus aurantiacus, Cyanobacteria, Dehalococcoides ethenogenes, Deinococcus radiodurans, Fervidobacterium islandicum, Helicobacter pylori, Mycoplasma spp., Pirellula marina, Propionibacterium freundenreichii, Simkania negevensis, Thermotoga maritima,* and *Ureaplasma urealyticum.*

EXAMPLE 11:

Methods for Eukaryotic rRNA Depletion from Mixed Human/*E. coli* Total RNA

These experiments were performed to demonstrate the depletion of 18S and 28S rRNAs from a mixture of prokaryotic and eukaryotic total RNA. Materials from Example 1 were used in the following experiments, except where noted.

RNA/Bridging Nucleic Acid Mixture Annealing

RNA (50 µg human total / 0.5 µg *E. coli* total) in 30 µl TE pH8.0 was added to 300 µl of binding buffer (The binding buffer in this example contains 0.02% Triton-X 100 which has been shown to reduce non-specific interactions between nucleic acids target and the capture nucleic acid). The bridging nucleic acid mixture consisted of d18S-3711, d18S-4238, d18S-5482, d28S-7979, d28S-11599 and d28S-12533 (each of these bridging oligonucleotides is at a final concentration of 3.33 µM). The bridging nucleic acid mixture (20 µl) was added to the RNA and the mixture was incubated at 70°C for 10 minutes and then shifted to 37°C for 1 hour.

Preparation of Capture Nucleic Acid

Capture nucleic acid (Oligo (dT) MagBeads, Seradyn) in storage buffer was mixed and 250 µl was removed to a separate tube. A magnetic stand was applied to the side of the tube to capture the magnetic beads and the supernatant was removed. The capture nucleic acid was equilibrated one time with distilled, deionized water (250 µl) and once with binding buffer (250 µl). The captured nucleic acid was captured again with a magnetic stand, and the binding buffer wash was removed. The magnetic beads were stored on ice until used in the next step (rRNA Capture).

rRNA Capture

Following the 1 hour annealing of RNA with the bridging nucleic acid mixture, the RNA-bridging oligonucleotide mixture was added to the capture nucleic acid (see Preparation of Capture Nucleic Acid), and the mixture was incubated at 37°C for 15 minutes. A magnetic stand was applied to the tube to capture the magnetic beads. The supernatant containing *E. coli* total RNA was removed to another tube and saved. An optional washing step was performed. The magnetic beads were washed with Wash Solution (100 µl) and captured again. The wash supernatant was removed and added to the original supernatant.

Fifteen minutes was found to be an adequate length of time for rRNA capture. Longer time periods can be used with no adverse effects. rRNA capture occurs rapidly, and capture times of 5 – 60 minutes have been used successfully in the methods of the invention.

Precipitating RNA

E. coli total RNA was precipitated by adding 1/10 volume of 3 M NaOAc (pH 5.5) and 3 volumes of 100% EtOH and incubating at –20°C for 60 minutes. The precipitated RNA was pelleted in a microfuge, washed with 70% EtOH, and resuspended in TE (pH 8.0).

Analysis of Purified RNA

Purified RNA was analyzed with the Caliper RNA 6000 LabChip kit on an Agilent Bioanalyzer. Purified RNA was compared with a control total RNA sample that was carried through the reaction as described above, except that the Bridging Nucleic Acid Mixture was omitted (FIG. 9). The percentage of a rRNA present in the sample is calculated from the area under the peak. The percentage removal of the 18S and 28S rRNAs was calculated as described in Example 2 for removal of 16S and 23S rRNAs.

Electropherograms of RNA from a control reaction and from an experimental reaction after ribosomal RNA depletion are shown in FIG. 9.

EXAMPLE 12:

Evaluation of efficacy with Mixed Prokaryotic and Eukaryotic rRNA Targets

The purpose of this experiment was to determine if one could, first remove the eukaryotic RNA and subsequently remove the prokaryotic rRNAs from a mixture of the two total RNAs. The materials and methods of Examples 1, 2 and 11 were used, except where noted. Depletion methods were verified using various mammalian samples, including rat liver total RNA, and both *E. coli* and *Bacillus subtilis* total RNA.

25 µg rat liver total RNA and 2 µg *E. coli* total RNA were mixed prior to the RNA enrichment procedure. The bridging oligonucleotides employed were: d16S-807, d16S-1092, d23S-1954, d23S-2511, d18S-3711, d18S-4238, d18S-5482, d28S-7979, d28S-11599, and d28S-12533.

The reactions began with procedures similar to Example 11, except for the following changes. 10 µl bridging nucleic acid mixture and 125 µl capture nucleic acid were used to remove the mammalian 18S and 28S rRNAs. Following the wash step, the wash solution and the unbound fraction containing the bacterial RNA were combined. The precipitation step was not performed. Instead the bacterial 16S and 23S rRNA was removed as in Example 2, with the following modifications. The bridging nucleic acid mixture (4 µl) containing d16S-807, d16S-1092, d23S-1954, and d23S-2511 was added directly to the combined wash and unbound

fractions containing the bacterial RNA. The remainder of the procedure for 16S and 23S rRNA followed the methods from Example 2.

Electropherograms of RNA from a control reaction with no bridging nucleic acids (FIG. 10A), from a reaction following 18S and 28S rRNA removal (FIG. 10B), and from a reaction following subsequent removal of 16S and 23S rRNA (FIG. 10C) are shown in FIG. 10.

EXAMPLE 13:
Use of Purified *E. coli* mRNA from Mixed Eukaryotic/Prokaryotic Samples in Gene Array Expression Analysis

mRNA was purified from total *E. coli* RNA (2 μg) in a background of human total RNA (25 μg) using the methods of the invention as described in Example 12 (16S, 23S, 18S and 28S rRNAs were all depleted from the sample). A control reaction was also performed in which the bridging nucleic acid mixtures were omitted from the reaction. Control total RNA (8.4 μg) and purified mRNA (1.0 μg) were added to 160 pmol random decamers in a final volume of 24.5 μl. These RNA amounts represent equal fractions of the control and purified RNA samples after the procedure is complete. The mixture was heated at 70°C for 10 minutes, then transferred to ice for 3 minutes. The following components were added to each reaction:

12 μl	cDNA 1st strand synthesis buffer (Invitrogen)
6 μl	0.1 M DTT
3 μl	10 mM dATP
3 μl	10 mM dGTP
3 μl	10 mM dTTP
5 μl	10 mCi/ml ^{33}P-dCTP (Perkin Elmer-NEN)
1 μl	RNase Inhibitor (cloned)

The reaction was then incubated at room temperature for 10 minutes, and the following component was added:

2 μl	Superscript II reverse transcriptase (Invitrogen) 200 U/μl

The reactions were incubated at 42°C for 120 minutes. Unincorporated nucleotides were removed from the reactions with a Qiaquick PCR™ cleanup column (Qiagen). RNA present in the cDNA probes was hydrolyzed by incubation at 65°C for 10 minutes in .05 N NaOH. The probes were subsequently neutralized with 0.05 M HCl. The labeled cDNAs (5×10^7 cpm/blot) were used to probe replicate portions of Panorama™ *E. coli* gene arrays, using hybridization buffers supplied by the array manufacturer (Sigma-Genosys). The arrays were washed and exposed to film. This example demonstrates a dramatic increase in hybridization signal (sensitivity) on gene arrays when labeled cDNA is prepared from enriched bacterial mRNA, purified according to the methods of the invention, rather than from the mixed prokaryotic and eukaryotic total RNA.

Example 14:

Evaluations of Efficacy with non-*E.coli* Prokaryotic Targets

The materials and methods of Examples 1 and 2 were employed in the Examples below except where noted. These experiments were performed to evaluate the efficacy of various bridging nucleic acids with different bacterial species.

Additional targeting regions for prokaryotic 16S and 23S rRNAs were designed. The targeting regions are shown, in the examples below, 3' of the bridging regions. Thus, the targeting region encompasses the remaining, non-bridging region of each molecule described below. SEQ ID NOs are provided for the targeting regions of the bridging nucleic acids provided below (*i.e.*, sequence of bridging regions not included in SEQ ID NO.). Furthermore, the oligos have been further designated with a suffix at the end of the oligo number. CY refers to cyanobacteria; P refers to pseudomonas; R refers to rhodobacter; and CH refers to campylobacter/helicobacter.

16S prokaryotic rRNA bridging oligonucleotides

d16S-1114P
5'-AAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTACGGGACTT-3' (SEQ ID NO:74)

d16S-1114R
5'-AAAAAAAAAAAAAAAAAAGGGTTGCGCTCGTTGCCGGACTT-3' (SEQ ID NO:75)

d16S-364
5'-AAAAAAAAAAAAAAAAAATCCCCACTGCTGCCTCCCGTAGG-3' (SEQ ID NO:76)

d16S-1087

5'-AAAAAAAAAAAAAAAAAACAACATCTCACGACACGAGCTGA-3' (SEQ ID NO:77)

d16S-364CY
5'-AAAAAAAAAAAAAAAAAATCCCCACTGCTGCCTCCCGTAGG-3' (SEQ ID NO:78)

d16S-534CY
5'-AAAAAAAAAAAAAAAAAATTACCGCGGCTGCTGGCACGGA-3' (SEQ ID NO:79)

d16S-928CY
5'-AAAAAAAAAAAAAAAAAACCCCGTCAATTCCTTTGAGTTTC-3' (SEQ ID NO:80)

d16S-1087CY
5'-AAAAAAAAAAAAAAAAAACAACATCTCACGACACGAGCTGA-3' (SEQ ID NO:81)

<u>23S prokaryotic rRNA bridging oligonucleotides</u>

d23S-479RCH
5'-AAAAAAAAAAAAAAAAAATTTCACCTTTCCCTCACGGTACT-3' (SEQ ID NO:82)

d23S-485
5'-AAAAAAAAAAAAAAAAAAGGTTCTTTTCACCTTTCCCTCGC-3' (SEQ ID NO:83)

d23S-518 CH
5'-AAAAAAAAAAAAAAAAAAATGGTTTCAGGTTCTATTTCACTC-3' (SEQ ID NO:84)

d23S-1954 CH
5'-AAAAAAAAAAAAAAAAAAATTTAACCGACAAGGAATTTCGC-3' (SEQ ID NO:85)

d23S-485CY
5'-AAAAAAAAAAAAAAAAAAGGTTCTTTTCACCTTTCCCTCGC-3' (SEQ ID NO:86)

The following results are from reactions that employed 10 μg of *Pseudomonas aeruginosa* total RNA, 40 pmol of 16S bridging nucleic acid and 40 pmol of 23S bridging nucleic acid, and 50 μl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	**% 16S Removed average of 2 reactions**	**% 23S Removed average of 2 reactions**
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-1954 (20 pmol), d23S-2511 (20 pmol)	90.1%	99.5%
d16S-807 (20 pmol), d16S-1114P (20 pmol) d23S-1954 (20 pmol), d23S-2511 (20 pmol)	97.0%	99.4%

The following results are from reactions that employed 10 µg of *Bacillus subtilis* total RNA, 40 pmol of 16S bridging nucleic acid and 40 pmol of 23S bridging nucleic acid, and 50 µl of capture nucleic acid described in Example 1.

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-1954 (20 pmol), d23S-2511 (20 pmol)	98.2%	96.2%

The following results are from reactions that employed 10 µg of *Campylobacter fetus* total RNA, 40 pmol of 16S bridging nucleic acid and 40 pmol of 23S bridging nucleic acid, and 50 µl of capture nucleic acid described in Example 1. The *Campylobacter fetus* 23S rRNA is processed into two fragments (FIG. 12).

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S (1260 nt fragment) Removed average of 2 reactions	% 23S (1667 nt fragment) Removed average of 2 reactions
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-479CH (20 pmol), d23S-2511 (20 pmol)	97.3%	97.7%	89.3%
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-518CH (20 pmol), d23S-2511 (20 pmol)	95.9%	96.7%	89.5%

The following results are from reactions that employed 10 µg of *Rhodobacter sphaeroides* total RNA, 40 pmol of 16S bridging nucleic acid and 40 pmol of 23S bridging nucleic acid, and 50 µl of capture nucleic acid, as described in Example 1 unless otherwise noted . The *Rhodobacter sphaeroides* 23S rRNA is processed into two fragments (FIG. 13). One fragment migrates with the 16S rRNA.

Bridging Nucleic Acid 16S/23S	% 16S + 23S fragment (1600 nt) Removed average of 2 reactions	% 23S Removed average of 2 reactions
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-479CH (20 pmol), d23S-2511 (20 pmol)	81.6%	96.8%
d16S-807 (20 pmol), d16S-1092 (20 pmol) d23S-518CH (20 pmol), d23S-2511 (20 pmol)	95.9%	96.7%
d16S-537 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-2511 (20 pmol)	89.3%	96.1%
d16S-537 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-1954 (20 pmol), d23S-2511 (20 pmol)	97.0%	96.4%
d16S-807 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-2511 (20 pmol)	83.1%	96.5%
d16S-807 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-1954 (20 pmol) d23S-2511 (20 pmol)	90.2%	95.9%
d16S-537 (20 pmol), d16S-807 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-1954 (20 pmol) d23S-2511 (20 pmol)	96.7%	94.9%
d16S-537 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-1954 (20 pmol) d23S-2511 (20 pmol)	95.0%	90.2%
d16S-537 (20 pmol), d16S-1114R(20 pmol) d23S-479CH (20 pmol), d23S-1954 (20 pmol) d23S-2511 (20 pmol)	95.2%	89.8%

The results demonstrate that bridging nucleic acids will function with various species. This example also demonstrates functionality based on sequence comparison, *i.e.*, that a bridging oligonucleotide will function with rRNAs in different organisms based on sequence identity between the oligonucleotide and the rRNA of the organism.

Evaluations of Efficacy with Cyanobacteria Targets

The materials and methods of Examples 1 and 2 were employed except where noted. These experiments were performed to evaluate the efficacy of various bridging nucleic acids with different bacterial species.

The following results are from reactions that employed 10 µg of *Anabaena spp.* total RNA, the indicated amounts of the bridging nucleic acids, and 50 µl of capture nucleic acid described in Example 1. The *Anabaena sp.* 23S rRNA and 23S rRNAs from other cyanobacteria may be processed into several fragments (FIG. 14).

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions		
		520 nt fragment	2090 nt fragment	2470 nt fragment
d16S-364CY (20 pmol), d16S-928CY (20 pmol)	94.3	NA	NA	NA
d16S-364CY (20 pmol), d16S-1087CY (20 pmol)	93.1	NA	NA	NA
d16S-928CY (20 pmol), d16S-1087CY (20 pmol)	93.7	NA	NA	NA
d23S-485 (20 pmol), d23S-1954 (20 pmol)	NA	98.6	94.1	99.5
d23S-485 (20 pmol), d23S-2511 (20 pmol)	NA	98.4	95.7	99.5
d16S-364CY (20 pmol), d16S-928CY (20 pmol) d23S-485, (20 pmol), d23S-1954 (20 pmol)	99.5	96.5	84.7	97.7
d16S-364CY (20 pmol), d16S-928CY (20 pmol) d23S-485, (20 pmol), d23S-2511 (20 pmol)	99.2	96.3	87.2	98.5
d16S-364CY (20 pmol), d16S-1087CY (20 pmol) d23S-485 (20 pmol), d23S-1954 (20 pmol)	99.7	98.6	88.3	98.3
d16S-364CY (20 pmol), d16S-1087CY (20 pmol) d23S-485 (20 pmol), d23S-2511 (20 pmol)	99.7	99.2	89.2	99.1
d16S-928CY (20 pmol), d16S-1087CY (20 pmol) d23S-485 (20 pmol), d23S-2511 (20 pmol)	96.9	97.6	86.8	98.9
d16S-364 (20 pmol), d16S-1087CY (20 pmol) d23S-485 (20 pmol), d23S-1954 (20 pmol)	99.2	98.2	88.1	97.7
d16S-364 (17.5 pmol), d16S-1087CY (17.5 pmol) d23S-485 (20 pmol), d23S-1954 (25 pmol)	99.8	98.8	88.4	98.5
d16S-364 (15 pmol), d16S-1087CY (15 pmol) d23S-485 (20 pmol), d23S-1954 (30 pmol)	99.8	99.1	90.6	98.3
d16S-364 (12.5 pmol), d16S-1087CY (12.5 pmol) d23S-485 (20 pmol), d23S-1954 (35 pmol)	99.9	98.7	92.5	98.9

Bridging Nucleic Acid 16S/23S	% 16S Removed average of 2 reactions	% 23S Removed average of 2 reactions		
		520 nt fragment	2090 nt fragment	2470 nt fragment
d16S-364 (10 pmol), d16S-1087CY (10 pmol) d23S-485 (20 pmol), d23S-1954 (40 pmol)	99.9	98.9	90.6	98.6

****

All of the compositions and methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

U.S. Application No. 09/854,412

U.S. Patent No. 4,486,539
U.S. Patent No. 4,563,419
U.S. Patent No. 4,659,774
U.S. Patent No. 4,682,195
U.S. Patent No. 4,683,202
U.S. Patent No. 4,751,177
U.S. Patent No. 4,816,571
U.S. Patent No. 4,868,105
U.S. Patent No. 4,894,325
U.S. Patent No. 4,959,463
U.S. Patent No. 5,124,246
U.S. Patent No. 5,141,813
U.S. Patent No. 5,200,314
U.S. Patent No. 5,214,136
U.S. Patent No. 5,216,141
U.S. Patent No. 5,223,618
U.S. Patent No. 5,264,566
U.S. Patent No. 5,273,882
U.S. Patent No. 5,288,609
U.S. Patent No. 5,378,825
U.S. Patent No. 5,412,087
U.S. Patent No. 5,428,148
U.S. Patent No. 5,432,272
U.S. Patent No. 5,445,934
U.S. Patent No. 5,446,137
U.S. Patent No. 5,457,025

U.S. Patent No. 5,466,786
U.S. Patent No. 5,470,967
U.S. Patent No. 5,500,356
U.S. Patent No. 5,539,082
U.S. Patent No. 5,554,744
U.S. Patent No. 5,574,146
U.S. Patent No. 5,589,335
U.S. Patent No. 5,602,240
U.S. Patent No. 5,602,244
U.S. Patent No. 5,610,289
U.S. Patent No. 5,614,617
U.S. Patent No. 5,623,070
U.S. Patent No. 5,645,897
U.S. Patent No. 5,652,099
U.S. Patent No. 5,670,663
U.S. Patent No. 5,672,697
U.S. Patent No. 5,681,947
U.S. Patent No. 5,700,922
U.S. Patent No. 5,702,896
U.S. Patent No. 5,708,154
U.S. Patent No. 5,709,629
U.S. Patent No. 5,714,324
U.S. Patent No. 5,714,331
U.S. Patent No. 5,714,606
U.S. Patent No. 5,719,262
U.S. Patent No. 5,723,597
U.S. Patent No. 5,736,336
U.S. Patent No. 5,744,305
U.S. Patent No. 5,759,777
U.S. Patent No. 5,763,167
U.S. Patent No. 5,766,855
U.S. Patent No. 5,773,571

U.S. Patent No. 5,777,092
U.S. Patent No. 5,786,461
U.S. Patent No. 5,792,847
U.S. Patent No. 5,858,988
U.S. Patent No. 5,859,221
U.S. Patent No. 5,872,232
U.S. Patent No. 5,886,165
U.S. Patent No. 5,891,625
U.S. Patent No. 5,897,783
U.S. Patent No. 5,908,845
U.S. Patent No. 5,945,525
U.S. Patent No. 6,001,983
U.S. Patent No. 6,013,440
U.S. Patent No. 6,037,120
U.S. Patent No. 6,060,246
U.S. Patent No. 6,090,548
U.S. Patent No. 6,110,678
U.S. Patent No. 6,140,496
U.S. Patent No. 6,203,978
U.S. Patent No. 6,221,581
U.S. Patent No. 6,228,580
U.S. Patent No. 6,309,823
U.S. Patent No. 6,316,193
U.S. Patent No. 6,322,971
U.S. Patent No. 6,324,479
U.S. Patent No. 6,329,140
U.S. Patent No. 6,329,209

EP 266,032
PCT/EP/01219
PCT/US00/29865

WO 01/32672

WO 86/05815

WO90/06045

WO 92/20702

The entire issue of *Current Opinion in Microbiology*, Volume 4, February 2001.

Amara *et al., Nucl. Acids Res.* 25:3465-3470, 1997.

Arfin *et al., J. Biol. Chem.* 275:29672-29684.

Ausubel *et al.*, In: *Current Protocols in Molecular Biology*, John, Wiley & Sons, Inc, New York, 1994.

Beaucage, *Methods Mol. Biol.* 20:33-61, 1993.

Chuang *et al., J. Bacteriol.* 175:2026-2036, 1993.

Coombes *et al., Infect. Immun.* 69:1420-1427, 2001.

Cornelis *et al., Curr. Opin. Microbiol.* 4:13-15, 2001.

Cummings *et al., Emerg. Inf. Dis.* 6:513-524, 2000.

DeRisi *et al., Nature Genetics* 14:457-460, 1996.

Detweller *et al., Proc. Natl. Acad. Sci. USA* 98:5850-5855, 2001.

Egholm *et al., Nature* 365(6446):566-568, 1993.

Feng *et al., Proc. Natl. Acad. Sci. USA* 97:6415-6420, 2000.

Fox, J.L. *et al., ASM News* 67:247-252, 2001.

Froehler *et al., Nucleic Acids Res.*, 14(13):5399-5407, 1986.

Gillam *et al., J. Biol. Chem.* 253(8):2532-9, 1978.

Gillam *et al., Gene* 8(1):99-106, 1979.

Gingeras *et al., ASM News* 66:463-469, 2000.

Graham *et al., Curr. Opin. Microbiol.* 4:65-70, 2001.

Graham *et al., Proc. Natl. Acad. Sci. USA* 96;11554-11559, 1999.

Ichikawa *et al., Proc. Natl. Acad. Sci. USA* 97:9659-9664, 2000.

Itakura *et al., J. Am. Chem. Soc.* 97(25):7327-32, 1975.

Kagnoff *et al., Curr. Opin. Microbiol.* 4:246-250, 2001.

Khorana, *Science* 203(4381):614-25, 1979.

Klug *et al., Methods Enzymol.* 152:316-325, 1987.

Koshkin *et al., Tetrahedron* 54:3607-3630, 1998.

Koshkin *et al., J. Am. Chem. Soc.* 120:13252-13253, 1998.

Kricka, *Nonisotopic DNA Probe Techniques*, Academic Press, San Diego, California, 1992.

Liang *et al., Methods Enzymol.* 254:304-321, 1995.

Lockhart *et al., Nature Biotech.* 14:1675, 1996.

Maskos *et al., Nuc. Acids. Res.* 20:1679-1684, 1992.

Neidhardt *et al., in Escherichia coli* and *Salmonella* (Neidhardt, FC, Ed.), Vol. 1, pp.13-16, ASM Press, Washington, DC, 1996.

Newton *et al., J Comput. Biol.* 8:37-52, 2001.

Pietu *et al., Genome Res.* 6:492, 1996.

Plum, *et al., Infect. Immun.* 62:476-483, 1994.

Rappuoli, R. *Proc. Natl. Acad. Sci. USA* 97:13467-13469, 2000.

Robinson *et al., Gene* 148:137-141, 1994.

Rosenberger *et al., J. Immunol.* 164:5894-5904, 2000.

Sambrook *et. al.,* In: *Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Sambrook *et al.*, In: *Molecular Cloning: A Laboratory Manual*, 3rd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY, 2001.

Schena *et al., Science* 270:467-470, 1995a.

Schena *et al., Proc. Natl. Acad. Sci. USA* 93:10539-11286, 1995b.

Shalon *et al., Genome Res.* 6:639-645, 1996.

Su *et al., Molec. Biotechnol.* 10:83-85, 1998.

Velculescu *et al., Science* 270:484-487, 1995.

Wahlestedt *et al., PNAS* 97:5633-5638, 2000.

Wei *et al., J. Bacteriol.* 183:545-556, 2001.

Wendisch, *et al., Anal. Biochem.* 290:205-213, 2001.

Wood *et al., Proc. Natl. Acad. Sci. USA.* 82:1585-1588, 1985.

Yoshida *et al., Nucl. Acids Res.* 29:683-692, 2001.

Zhao *et al., Gene* 156:207, 1995.

CLAIMS

1. A method for depleting or isolating a targeted nucleic acid from a sample comprising:
   a) incubating the sample with a first bridging oligonucleotide comprising (1) at least one bridging region comprising at least 5 nucleic acid residues and (2) at least one targeting region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the first targeting region and the targeted nucleic acid;
   b) incubating the first bridging oligonucleotide with a capture oligonucleotide comprising a nonreacting structure and a capture region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the bridging region and the capture region; and
   c) isolating the targeted nucleic acid from the remainder of the sample.

2. The method of claim 1 wherein the targeted nucleic acid is rRNA.

3. The method of claim 2, wherein the rRNA is prokaryotic 16S, prokaryotic 23S, eukaryotic 17S or 18S, or eukaryotic 28S rRNA.

4. The method of claim 3, wherein the rRNA comprises the sequence of SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, or SEQ ID NO:86.

5. The method of claim 1, wherein the sample comprises eukaryotic nucleic acid.

6. The method of claim 1, wherein the sample comprises prokaryotic nucleic acid.

7. The method of claim 6, wherein the prokaryotic nucleic acid is from a gram positive bacterium.

8. The method of claim 6, wherein the prokaryotic nucleic acid is from a gram negative bacterium.

9. The method of claim 1, wherein the bridging region, targeting region, or capture region comprises at least 10 nucleic acid residues.

10. The method of claim 9, wherein the bridging region, targeting region, or capture region comprises at least 15 nucleic acid residues.

11. The method of claim 10, wherein the bridging region, targeting region, or capture region comprises at least 20 nucleic acid residues.

12. The method of claim 1, wherein the bridging region or the capture region is polypurine or polypyrimidine.

13. The method of claim 12, wherein the bridging region is polypurine and the capture region is polypyrimidine.

14. The method of claim 1, further comprising incubating the sample with a second bridging oligonucleotide comprising (1) at least one bridging region comprising at least 5 nucleic acid residues and (2) at least one targeting region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the targeting region of the second bridging oligonucleotide and the targeted nucleic acid.

15. The method of claim 14, wherein the targeting region of the first bridging oligonucleotide is complementary to the sequence of a targeted nucleic acid and the targeting region of the second bridging oligonucleotide is complementary to a different sequence of a targeted nucleic acid.

16. The method of claim 15, wherein the targeting region of the first bridging oligonucleotide and the targeting region of the second bridging oligonucleotide are complementary to the same targeted nucleic acid.

17. The method of claim 15, wherein the targeting region of the first bridging oligonucleotide and the targeting region of the second bridging oligonucleotide are complementary to different targeted nucleic acids.

18. The method of claim 17, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of the largest rRNA molecule and the targeting region of the second bridging oligonucleotide is complementary to a sequence of the second largest rRNA molecule in the sample.

19. The method of claim 14, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 100 and 5000 residues of the 5' or 3' end of the targeted nucleic acid.

20. The method of claim 19, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 150 and 4000 residues of the 5' or 3' end of the targeted nucleic acid.

21. The method of claim 20, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 200 and 3000 residues of the 5' or 3' end of the targeted nucleic acid.

22. The method of claim 21, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 250 and 2000 residues of the 5' or 3' end of the targeted nucleic acid.

23. The method of claim 22, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 300 and 1500 residues of the 5' or 3' end of the targeted nucleic acid.

24. The method of claim 23, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 350 and 1000 residues of the 5' or 3' end of the targeted nucleic acid.

25. The method of claim 24, wherein targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 400 and 900 residues of the 5' or 3' end of the targeted nucleic acid.

26. The method of claim 25, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 450 and 800 residues of the 5' or 3' end of the targeted nucleic acid.

27. The method of claim 26, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence located between 500 and 700 residues of the 5' or 3' end of the targeted nucleic acid.

28. The method of claim 14, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence at the 3' or 5' end of the targeted nucleic acid.

29. The method of claim 14, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence not within 100 residues from the 3'or 5' end of the targeted nucleic acid.

30. The method of claim 14, wherein targeting region of the first or second bridging oligonucleotide hybridizes to a sequence not within 200 residues from the 3'or 5' end of the targeted nucleic acid.

31. The method of claim 14, wherein the targeting region of the first or second bridging oligonucleotide hybridizes to a sequence not within 400 residues from the 3'or 5' ends of the targeted nucleic acid.

32. The method of claim 14, wherein the targeting region of the first or second bridging oligonucleotide comprises SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, or SEQ ID NO:86.

33. The method of claim 1, wherein the bridging oligonucleotide comprises a second targeting region comprising at least 5 nucleic acid residues complementary to a different sequence than the sequence to which the first targeting region is complementary.

34. The method of claim 33, wherein the first targeting region is complementary to a different targeting nucleic acid than the second targeting region is.

35. The method of claim 1, wherein the first bridging oligonucleotide comprises two bridging regions.

36. The method of claim 1, wherein the bridging oligonucleotide or the capture oligonucleotide is RNA, DNA, LNA, iso-bases, or a peptide nucleic acid.

37. The method of claim 1, further comprising washing the capture oligonucleotide after incubation with the sample and the bridging oligonucleotide.

38. The method of claim 1, wherein a) and b) are performed at the same temperature.

39. The method of claim 1, wherein a) and b) are performed at a different temperature.

40. The method of claim 38, wherein a) and b) are performed at the same time.

41. The method of claim 1, wherein the nonreacting structure comprises a bead comprising plastic, glass, teflon, silica, a magnet, cellulose, latex, polystyrene, nylon, cellulose, nitrocellulose, polymethacrylate, polyvinylchloride, or styrene-divinylbenzene

42. The method of claim 41, wherein isolating the targeted nucleic acid away from the sample comprises exposing the sample with the capture oligonucleotide to a magnetic field.

43. The method of claim 1, wherein the nonreacting structure is cellulose.

44. The method of claim 1, wherein the nonreacting structure is biotin.

45. The method of claim 44, wherein isolating the targeted nucleic acid comprises incubating the sample with streptavidin or avidin.

46. The method of claim 1, wherein the sample, capture oligonucleotide, and bridging oligonucleotide are incubated in a buffer comprising TMAC or TEAC.

47. The method of claim 1, further comprising:

d) discarding the portion of the sample that hybridizes to the capture oligonucleotide.

48. The method of claim 2, further comprising:

d) discarding the targeted rRNA nucleic acid; and

e) producing cDNA using mRNA in the remainder of the sample.

49. The method of claim 2, further comprising:

d) amplifying nucleic acids in the remainder of the sample, wherein the remainder of the sample is enriched for mRNA.

50. The method of claim 49., further comprising:

e) using the amplified nucleic acids to probe a nucleic acid array.

51. The method of claim 48, further comprising:

f) attaching the cDNA to a solid support, wherein a nucleic acid array is created.

52. The method of claim 51, wherein the solid support is plastic, glass, or nylon.

53. The method of claim 52, wherein the solid support is a plate.

54. The method of claim 53, wherein the plate is a multiple-well plate.

55. The method of claim 48, further comprising:

f) contacting a nucleic acid array with the cDNA.

56. The method of claim 1, further comprising incubating the sample with a second bridging oligonucleotide comprising (1) at least one bridging region comprising at least 5 nucleic acid residues and (2) at least one targeting region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the targeting region of the second bridging oligonucleotide and a second targeted nucleic acid.

57. The method of claim 56, further comprising incubating the sample with a third bridging oligonucleotide comprising (1) at least one bridging region comprising at least 5 nucleic acid residues and (2) at least one targeting region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the targeting region of the third bridging oligonucleotide and a third targeted nucleic acid.

58. The method of claim 57, further comprising incubating the sample with a fourth bridging oligonucleotide comprising (1) at least one bridging region comprising at least 5 nucleic acid residues and (2) at least one targeting region comprising at least 5 nucleic acid residues, under conditions allowing hybridization between the targeting region of the fourth bridging oligonucleotide and a fourth targeted nucleic acid.

59. The method of claim 56, wherein prokaryotic and eukaryotic rRNAs are targeted nucleic acids.

60. A method for depleting rRNA from a sample comprising:

a) incubating the sample with at least a first (1) bridging oligonucleotide comprising a bridging region comprising a poly-purine region of at least 5 residues and a targeting region comprising at least 5 contiguous nucleic acid residues complementary to a sequence of an rRNA molecule and a (2) capture oligonucleotide comprising a magnetic bead and a capture region comprising a poly-pyrimidine region of at least 5 residues, under conditions to allow hybridization between the bridging oligonucleotide and the capture oligonucleotide and the bridging oligonucleotide and the rRNA;

b) incubating the sample with a magnetic bead; and

c) isolating the magnetic bead.

61. A kit, in a suitable container means, comprising:

a) a capture oligonucleotide comprising a capture region and a magnetic bead; and

b) at least a first bridging oligonucleotide comprising (1) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and a (2) at least one targeting region comprising 10 contiguous nucleic acids complementary to a sequence of an rRNA.

62. The kit of claim 61, wherein the first bridging oligonucleotide comprises a second targeting region.

63. The kit of claim 62, wherein the first and second targeting regions have the same nucleic acid sequence.

64. The kit of claim 62, wherein the first and second targeting regions have different nucleic acid sequences.

65. The kit of claim 64, wherein the first targeting region is complementary to a sequence of an eukaryotic rRNA and the second targeting region is complementary to a sequence of a prokaryotic rRNA.

66. The kit of claim 64, wherein the first targeting region is complementary to a sequence of an eukaryotic rRNA and the second targeting region is complementary to a sequence of a different eukaryotic rRNA than the first targeting region.

67. The kit of claim 64, wherein the first targeting region is complementary to a sequence of a prokaryotic rRNA and the second targeting region is complementary to a sequence of a different prokaryotic rRNA than the first targeting region.

68. The kit of claim 61, further comprising a second bridging oligonucleotide comprising (1) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and a (2) at least one targeting region comprising 10 contiguous nucleic acids complementary to a sequence of an rRNA.

69. The kit of claim 68, wherein the targeting region of the second bridging oligonucleotide is complementary to a sequence of the same rRNA as the first targeting region.

70. The kit of claim 68, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of the largest rRNA and the targeting region of the second bridging oligonucleotide is complementary to a sequence of the second largest rRNA in the sample.

71. The kit of claim 68, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of an eukaryotic rRNA and the targeting region of the bridging oligonucleotide is complementary to a sequence of a prokaryotic rRNA.

72. The kit of claim 70, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of an eukaryotic 28S rRNA and the targeting region of the second bridging oligonucleotide is complementary to a sequence of a eukaryotic 17S or 18S rRNA.

73. The kit of claim 70, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of a prokaryotic 23S rRNA and the targeting region of the second bridging oligonucleotide is complementary to a sequence of a prokaryotic 16S rRNA.

74. The kit of claim 70, wherein the targeting region of the first bridging oligonucleotide is complementary to a sequence of an eukaryotic 28S rRNA and the targeting region of the second bridging oligonucleotide is complementary to a sequence of a prokaryotic 23S rRNA.

75. The kit of claim 68, further comprising a third bridging oligonucleotide comprising (1) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and a (2) at least one targeting region comprising 10 contiguous nucleic acids complementary to a sequence of an rRNA.

76. The kit of claim 75, wherein the targeting region of the third bridging oligonucleotide is complementary to a sequence of a prokaryotic 23S rRNA.

77. The kit of claim 75, wherein the targeting region of the third bridging oligonucleotide is complementary to a sequence of a eukaryotic 18S rRNA.

78. The kit of claim 75, further comprising a fourth bridging oligonucleotide comprising (1) at least one bridging region complementary to all or part of the capture region of the capture oligonucleotide and a (2) at least one targeting region comprising 10 contiguous nucleic acids complementary to a sequence of an rRNA.

79. The kit of claim 78, wherein (i) the targeting region of the first bridging oligonucleotide is complementary to a sequence of a prokaryotic 16S rRNA, (ii) the targeting region of the second bridging oligonucleotide is complementary to a sequence of a prokaryotic 23S rRNA, (iii) the targeting region of the third bridging oligonucleotide is complementary to a sequence of a eukaryotic 18S rRNA, and (iv) the targeting region of the fourth bridging oligonucleotide is complementary to a sequence of a eukaryotic 28S rRNA,

80. The kit of claim 61, wherein the first targeting region of the bridging oligonucleotide comprises SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, or SEQ ID NO:86.

81. The kit of claim 61, further comprising a buffer comprising TMAC or TEAC.

82. The kit of claim 61, further comprising a magnetic stand.

83. The kit of claim 61, further comprising:

c) a solid support for preparing a nucleic acid array.

84. A bridging oligonucleotide comprising a (1) bridging region comprising a polypyrimidine or polypurine stretch and a (2) targeting region comprising at least 10 contiguous nucleic acid residues complementary to a sequence of an rRNA.

85. The oligonucleotide of claim 84, wherein the rRNA is eukaryotic.

86. The oligonucleotide of claim 85, wherein the rRNA is the 28S rRNA.

87. The oligonucleotide of claim 84, wherein the rRNA is prokaryotic.

88. The oligonucleotide of claim 87, wherein the rRNA is the 23S rRNA.

89. A method for depleting or isolating a targeted rRNA from a sample comprising:

a) obtaining the kit of claim 61;

b) incubating the sample with the bridging oligonucleotide under conditions allowing hybridization between the targeting region and the targeted rRNA;

c) incubating the bridging oligonucleotide with the capture oligonucleotide under conditions allowing hybridization between the bridging region and the capture region; and

d) isolating the targeted rRNA from the remainder of the sample by incubating the sample with a magnetic field.

90. The method of claim 89, further comprising:

e) obtaining the remainder of the sample enriched for mRNA;

f) preparing cDNA from the mRNA.

91. The method of claim 90, further comprising:

g) constructing a nucleic acid array with the cDNA.

92. The method of claim 89, wherein the mRNA or prepared cDNA is amplified.

93. The method of claim 92, wherein the cDNA is used to probe a nucleic acid array.

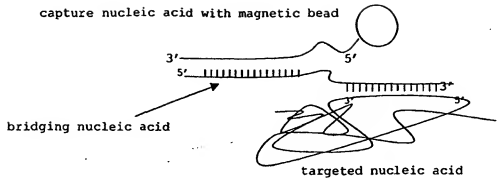

capture nucleic acid with magnetic bead
3′
5′
5′
3′
3′
5′
bridging nucleic acid
targeted nucleic acid

FIG. 1

Alignment Report of Gram +& - 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

FIG. 2A-1

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

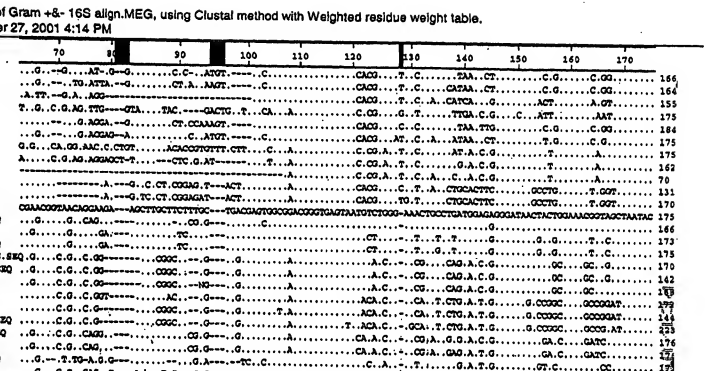


FIG. 2A-2

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

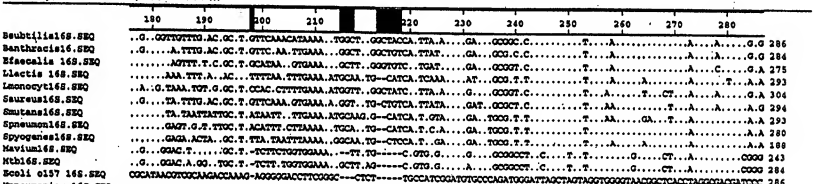


FIG. 2A-3

Allgnment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

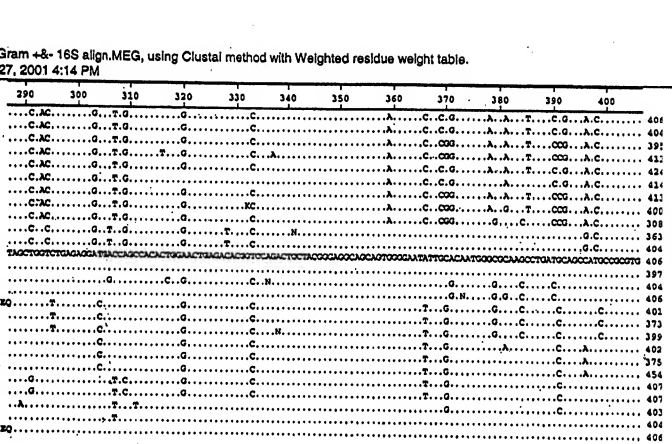


FIG. 2A-4

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

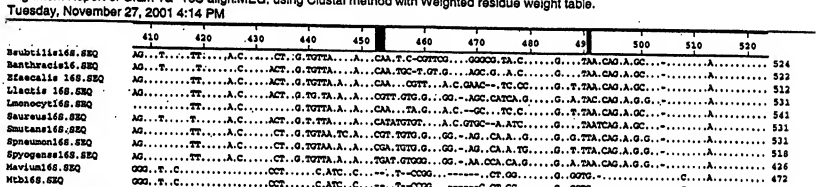


FIG. 2A-5

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

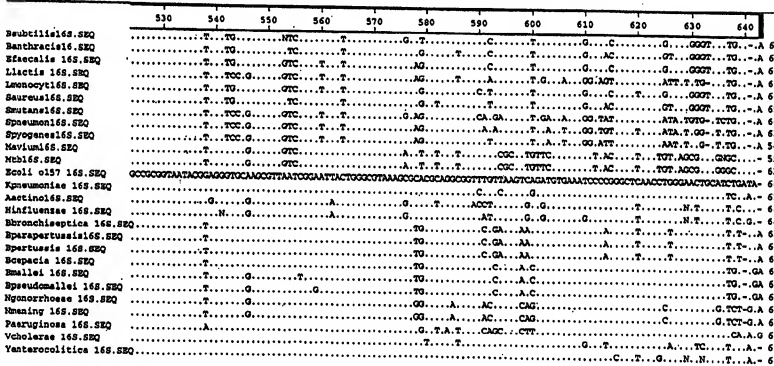


FIG. 2A-6

Allgnment Report of Gram +& 16S allgn.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

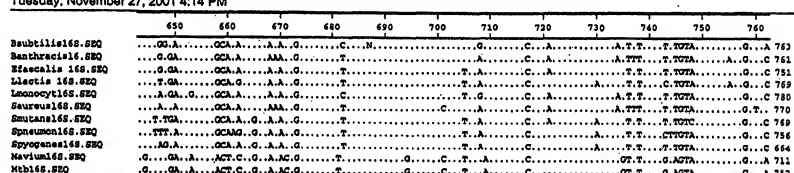


FIG. 2A-7

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

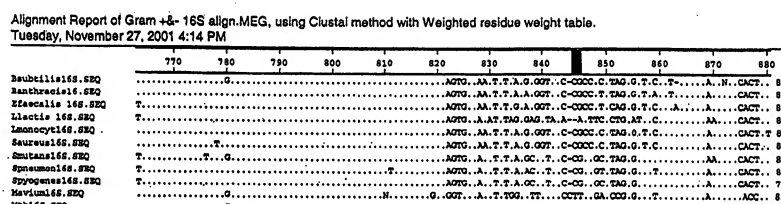


FIG. 2A-8

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

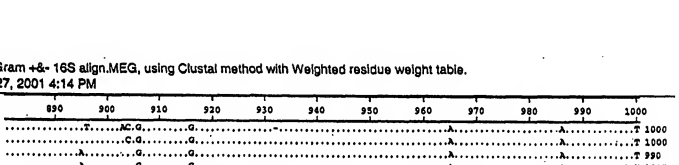


FIG. 2A-9

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

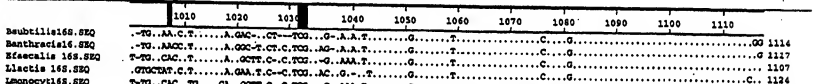


FIG. 2A-10

Alignment Report of Gram +& - 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

FIG. 2A-11

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

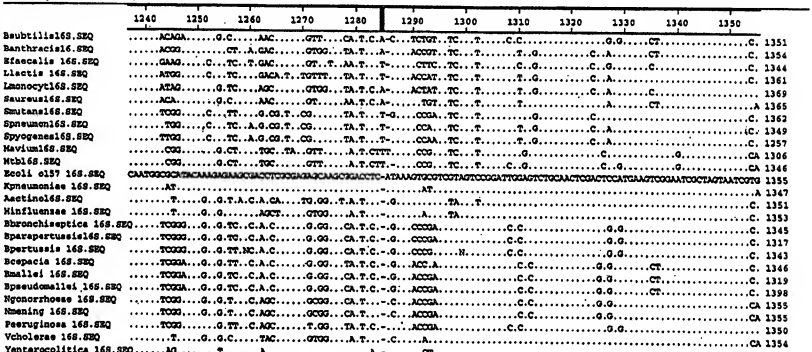


FIG. 2A-12

Alignment Report of Gram +&- 16S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

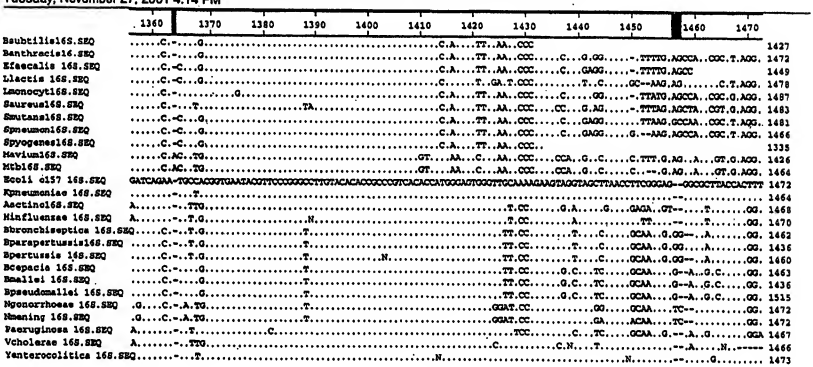
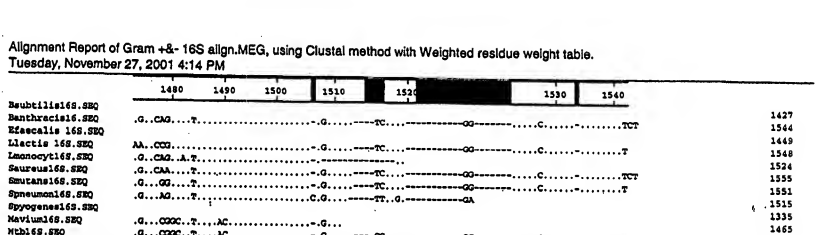


FIG. 2A-13

Allgnment Report of Gram +&- 16S allgn.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM



Decoration 'Decoration #1': Hide (as '.') residues that match Ecoli o157 16S.SEQ exactly.

FIG. 2A-14

Alignment Report of Gram +& 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

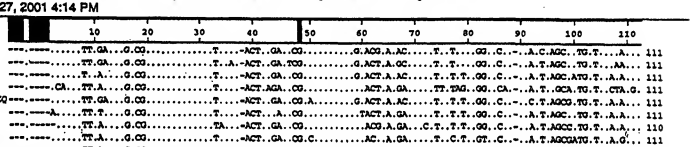


FIG. 2B-1

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

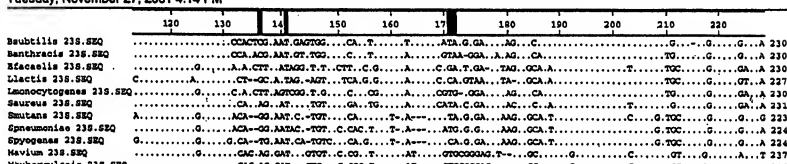


FIG. 2B-2

Allignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

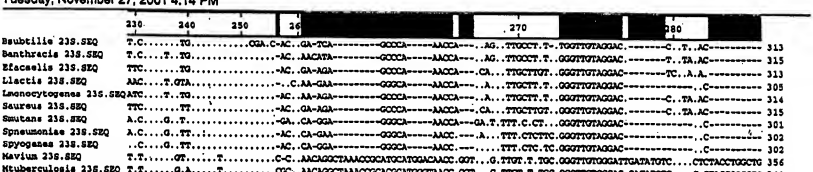


FIG. 2B-3

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

FIG. 2B-4

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

FIG. 2B-5

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

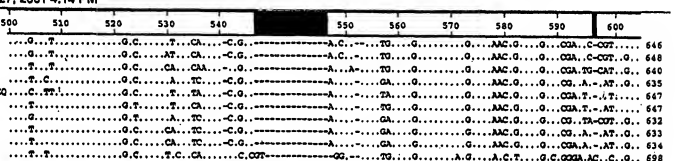


FIG. 2B-6

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

FIG. 2B-7

Allgnment Report of Gram +&- 23S allgn.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

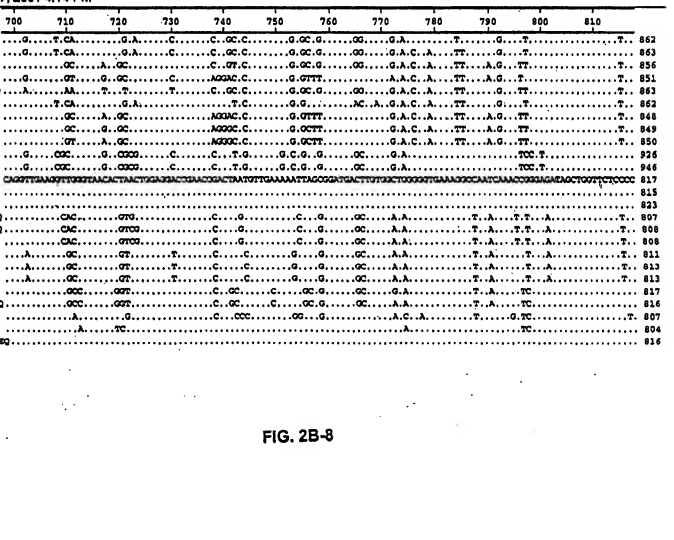


FIG. 2B-8

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

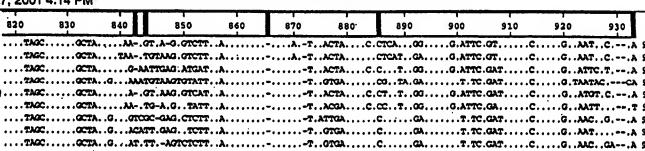


FIG. 2B-9

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

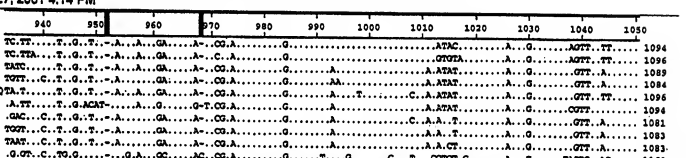


FIG. 2B-10

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

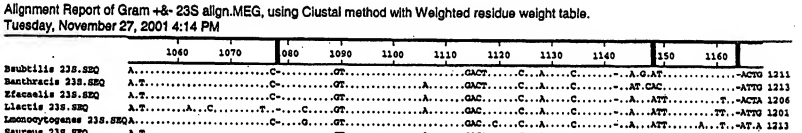


FIG. 2B-11

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

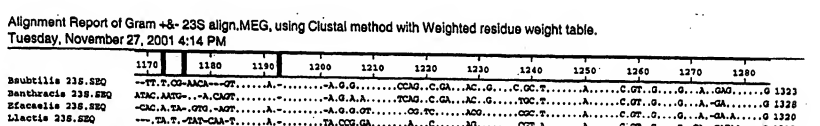


FIG. 2B-12

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

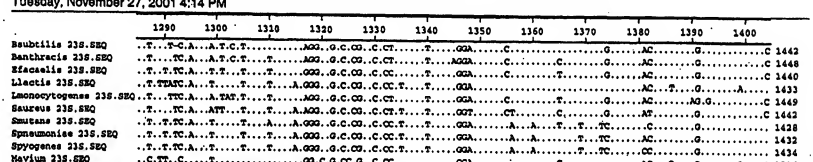


FIG. 2B-13

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

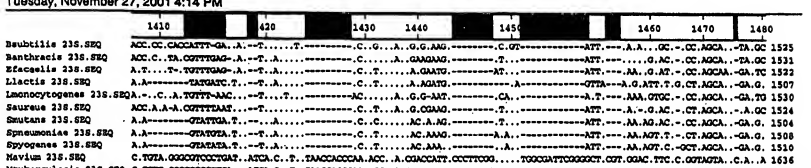


FIG. 2B-14

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

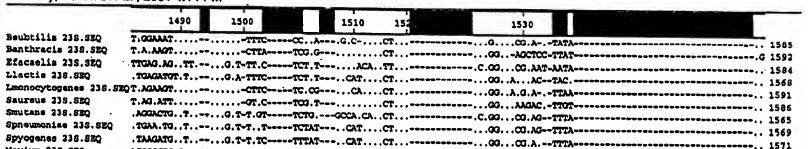


FIG. 2B-15

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

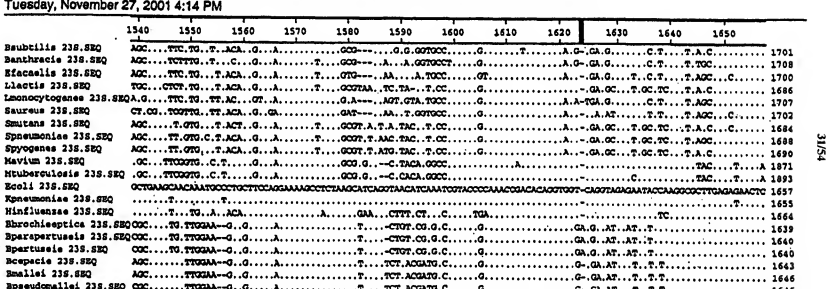


FIG. 2B-16

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

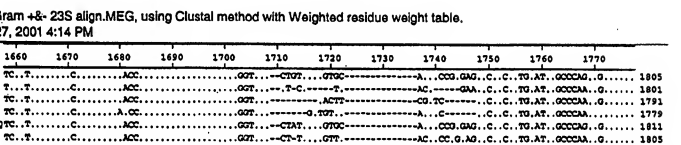


FIG. 2B-17

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

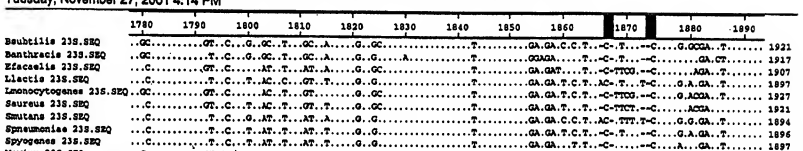


FIG. 2B-18

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

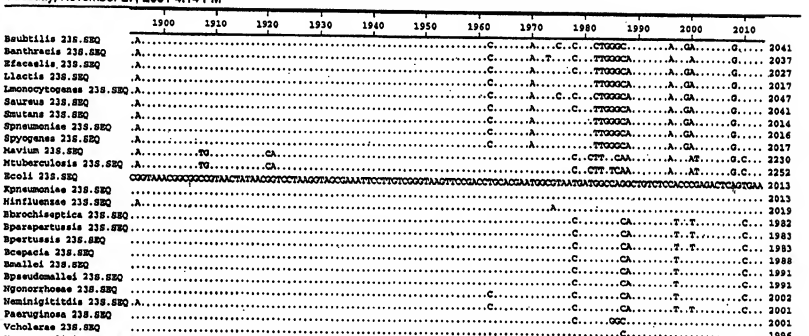


FIG. 2B-19

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

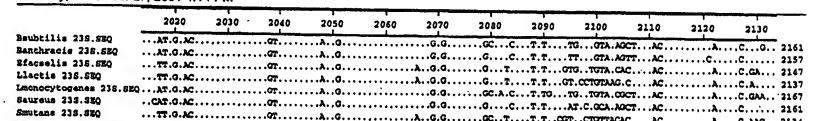


FIG. 2B-20

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

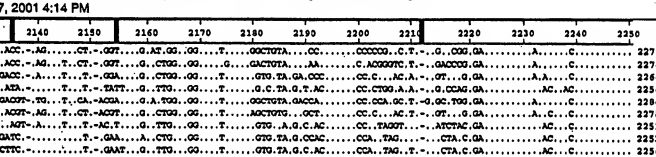


FIG. 2B-21

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:14 PM

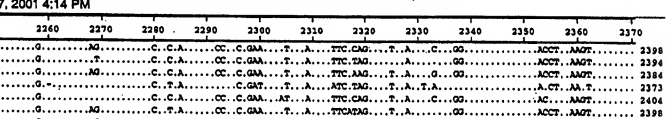


FIG. 2B-22

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:15 PM

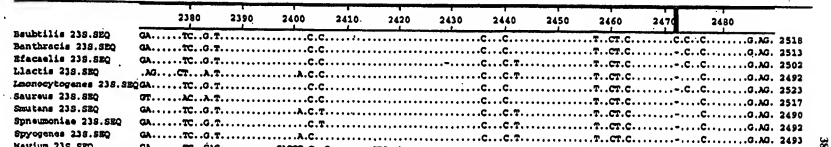


FIG. 2B-23

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:15 PM

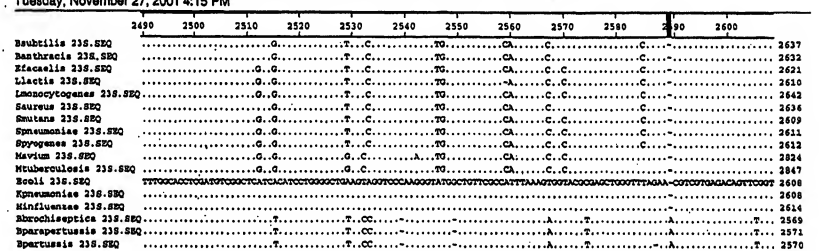


FIG. 2B-24

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:15 PM

```
                2610        2620        2630          40        2650        2660        2670        2680        2690        2700        2710        2720
Bsubtilis 23S.SEQ  .....C.T..C.........A.TT....--A..A....TC.T.....................GA.......C.G.......A.CA.....TC.....G....TC..TG.........TG 2755
Banthracis 23S.SEQ .....C.T.......TA...A.TT....--A..A....TC.T.....................GA.......C.G.......A.CA.....TC.....G....TA..TG.........TG 2750
Efaecalis 23S.SEQ  .....C.T..C.....T....A.TT....--A..A....TC.T.....................GA.....TT.C.G.......A.CA.....TC.....G....T...TG.........TG 2739
Llactis 23S.SEQ    .....C.T..C.....TA...TA.TT....--A..AT...TC.T.....................GA.....TT.C.G.......A.CA.....TCC....GGA....G...TG.A.....TG 2728
Lmonocytogenes 23S.SEQ.....C.T..C.....A...A.TT....--A..A....TC.T.....................GA.....A..C.G.......A.CA.....TCC....GGA....TC..TG.........TG 2760
Saureus 23S.SEQ    .....C.T.......TA...A.TT....--A..A....TC.T.....................GA.....AT.C.T.......A.CA.....G.....C....TA..TG.........TG 2754
Smutans 23S.SEQ    .....C.T..C....AA...A.TT....--A..AT.....................A.....TT.C.G.......A.CA.....TC.....GA...TC..TG.........G 2727
Spneumoniae 23S.SEQ .....C.T..C....TA...A.TT....--A..AT.....................A.....TT.C.G.......A.CA.....T.....A....TC..TG.........TG 2729
Spyogenes 23S.SEQ  .....C.T..C....TA...A.TT....--A..AT.....................A.....TT.C.G.......A.CA.....T.....A....TC..TG.........TG 2730
Mavium 23S.SEQ     .T.....C....C.C...TCA..A.CT....--.AAAC...TC.....................GAC.....A.C.T.....A.A.CA.....CCA..GG.....G..TG.A....C.CG 2942
Mtuberculosis 23S.SEQ .T.....C....C.C...TCA..A.CT....--.AAAC...TC.....................GAC.....A.C.T.....CA.CA.....CC....GG....C..TG.A....C.CG 2965
Ecoli 23S.SEQ      CCCTATCTGCGGTGGGCGCTGGAGAACTAG--GGGGGCTGCTCCTAGTACGAGAGACCGGAGTGGACGCATCACTGGTGTTCGGGTTGTCATGCCAATGGCACTGCCCGGTAGCTAAA 2726
Kpneumoniae 23S.SEQ ..................T....--.....................................................C.....GTC....GAC...T....G..... 2726
Hinfluenzae 23S.SEQ ..................TA...TG.T...T--T.....................................................C.....GTC....GAC...T....G..... 2732
Bbronchiseptica 23S.SEQ...............T....T.CT...ACA..A.C.....................................T.C.T.....A.C.....................T....G.........G 2689
Bparapertussis 23S.SEQ..............T....T.CT...ACA..A.C.....................................T.C.T.....A.C.....................T....G.........G 2691
Bpertussis 23S.SEQ  ................T....T.CT...ACA..A.C.....................................T.C.T.....A.C.....................T....G.........G 2690
Bcepacia 23S.SEQ    ................T....T.TT...A--.....................................................A.C.T.....A.C.....C....G.....TC...G.....TG 2701
Bmallei 23S.SEQ     ................T....AGTT...A--.....................................................A.C.T.....A.C.....G.C....G.C...TC...G.....TG 2704
Bpseudomallei 23S.SEQ...............T....AGTT...A--.....................................................A.C.T.....A.C.....G.C....G.C...TC...G.....TG 2704
Ngonorrhoeae 23S.SEQ ..........A....-..T....AGTT...C--..-.....................................................A.C.T.....A.C.....A.C....G.T...TA...G.....G 2713
Neminigititdis 23S.SEQ..........A.....T....AGTT...C--.....................................................A.C.T.....A.C.....A.C....G.T...TA...G.....G 2714
Paeruginosa 23S.SEQ ..............A.T.T....ST....--A.....................................................A.C.T.....C.....C....G.....T...G.....TG 2714
Vcholerae 23S.SEQ  ................T....AG.T...A--.....................................................A.C.T.....GTC....GAC...T....G.....G 2709
Yenterocolitica 23S.SEQ.............Y....RR.Y....--R.....................................................M.....Y.....T....G.....G 2729
```

FIG. 2B-25

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:15 PM

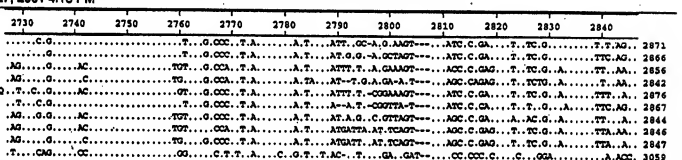


FIG. 2B-26

Alignment Report of Gram +&- 23S align.MEG, using Clustal method with Weighted residue weight table.
Tuesday, November 27, 2001 4:15 PM

```
                    2850       2860       2870       2880       2890       2900
Beubtilis 23S.SEQ   .G...T.TG..A.CA.A.G....G..A.A.....CG.T..-AG.A.....A-T   2927
Banthracis 23S.SEQ  .G....ATG.T..CAT..G.....G..GAA.....AG.T..-AG.A.....A-T   2922
Efaecalis 23S.SEQ   .G..CCT..T...AGT.G....GG...AA.....CGT..-AG.A.....A-A   2912
Llactis 23S.SEQ     .G...A.TT...GA.T.G....GG...A.....CGT..-AG.A...T..A-A   2898
Lmonocytogenes 23S.SEQ .G...T.T.....A.A.G....GG..AAA.....CG.T..-AG.A.....A-A   2932
Saureus 23S.SEQ     .G....ATG.T..CAT..G.....G..GAA.....CG.T...A..A.....T.A-A   2923
Smutans 23S.SEQ     .G.....TT.T..GA...GA...GG...AA.....CGT..-AG.A.....T.A-A   2900
Spneumoniae 23S.SEQ .G...T.TG....CA.A.GT...GG..TAA.....AGT..-AG.A....T..A-A   2902
Spyogenes 23S.SEQ   .....T.T.....A.A.GT...GG..TAA.....AGT..-AG.A.....T..     2901
Mavium 23S.SEQ      .G....T...TA...A..GC...GG...T..C....CTGG...-A.A..          3107
Mtuberculosis 23S.SEQ .G....T...TA...G..GT..GG...T...G....CCGG...-A.AA....-.AACA   3138
Ecoli 23S.SEQ       GTAAGGCAGCGGATGGTTGAGCTAACGGTACTAATGAACCGTGAGGCTTAACCTT   2903
Kpneumoniae 23S.SEQ ....................................................   2903
Hinfluenzae 23S.SEQ .....T.AT.T..GT.A.....AA.....TGC..A.....              2897
Bbrochiseptica 23S.SEQ .G.......TA.....A.....A.....TGC.............T.C.   2865
Bparapertussis 23S.SEQ .G.......TA.....A.....A.....TGC.............G.T.C.   2865
Bpertussis 23S.SEQ  .G.......TA.....A.....A.....TGC.............G.T.C.   2864
Bcepacia 23S.SEQ    .........TA.....C.....T.A.....TGC...A.....G.T.C.   2878
Bmallei 23S.SEQ     .........TA.....C.....A.....TGC...AC.....G.T.C.A   2882
Bpseudomallei 23S.SEQ .........TA.....C.....A.....TGC...AC.....G.T.C.A   2882
Ngonorrhoeae 23S.SEQ .G......G.TA.C...GA.....CA.....TGC.............G..TC.   2890
Neninigitidis 23S.SEQ .G......G.TA.C...GA.....CA.....TGCT..........G..TC.   2891
Paeruginosa 23S.SEQ .......TT.T..G.....A.....TGC.............G...A.        2891
Vcholerae 23S.SEQ   .......TT.T..G.....T.....TGC.............A.            2886
Yenterocolitica 23S.SEQ.....GAGTTG.....................................   2906

Decoration 'Decoration #1': Hide (as '.') residues that match Ecoli 23S.SEQ exactly.
```

FIG. 9B-27

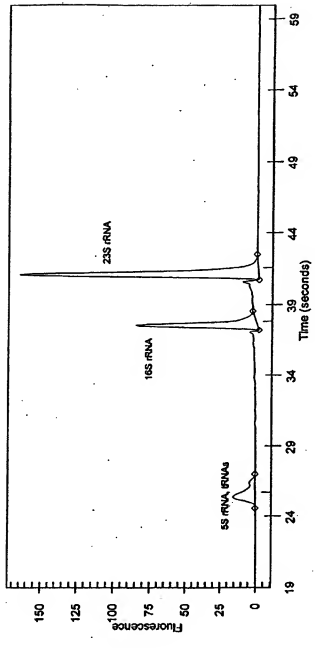

23S rRNA
16S rRNA
5S rRNA, tRNAs
Fluorescence
150
125
100
75
50
25
0
19
24
29
34
39
44
49
54
59
Time (seconds)

FIG. 3

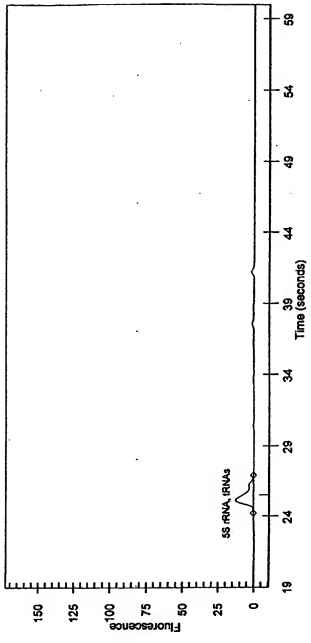

Fluorescence
150
125
100
75
50
25
0
5S rRNA, tRNAs
19
24
29
34
39
44
49
54
59
Time (seconds)

FIG. 4

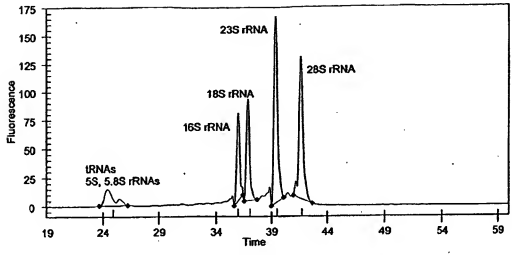

175
150
125
100
75
50
25
0
Fluorescence
23S rRNA
28S rRNA
18S rRNA
16S rRNA
tRNAs
5S, 5.8S rRNAs
19
24
29
34
39
44
49
54
59
Time

FIG. 5A

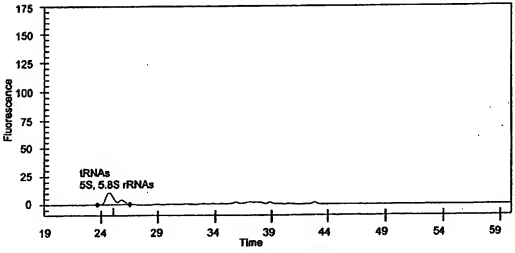

175
150
125
100
75
50
25
0
Fluorescence
tRNAs
5S, 5.8S rRNAs
19
24
29
34
39
44
49
54
59
Time

FIG. 5B

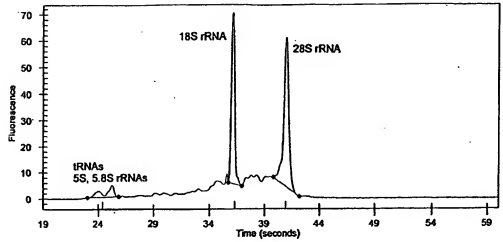

18S rRNA
28S rRNA
tRNAs
5S, 5.8S rRNAs
Fluorescence
Time (seconds)

FIG. 6A

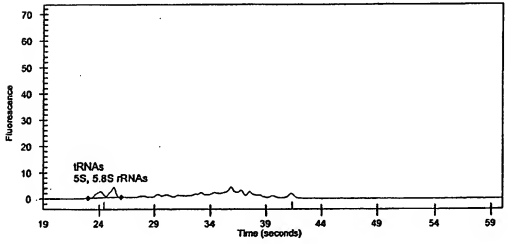

tRNAs
5S, 5.8S rRNAs
Fluorescence
Time (seconds)

FIG. 6B

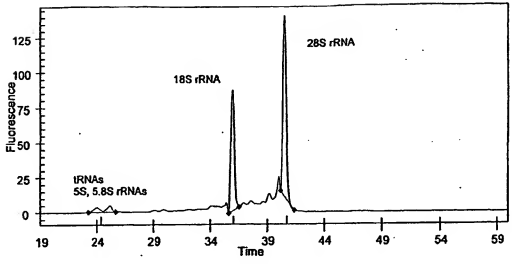

28S rRNA
18S rRNA
tRNAs
5S, 5.8S rRNAs
Fluorescence
125
100
75
50
25
0
19
24
29
34
39
44
49
54
59
Time

FIG. 7A

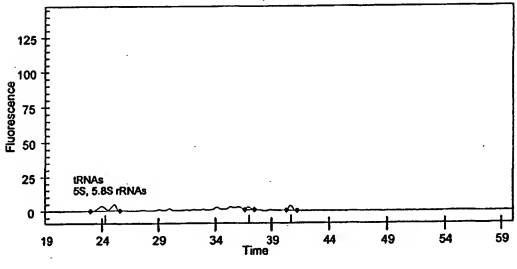

tRNAs
5S, 5.8S rRNAs
Fluorescence
125
100
75
50
25
0
19
24
29
34
39
44
49
54
59
Time

FIG. 7B

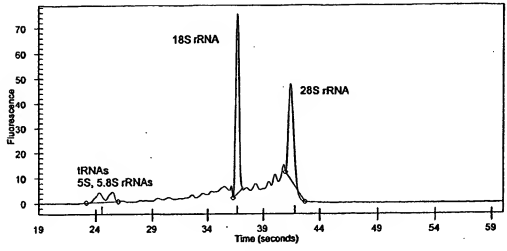

18S rRNA
28S rRNA
tRNAs
5S, 5.8S rRNAs
Fluorescence
Time (seconds)
70
60
50
40
30
20
10
0
19
24
29
34
39
44
49
54
59

FIG. 8A

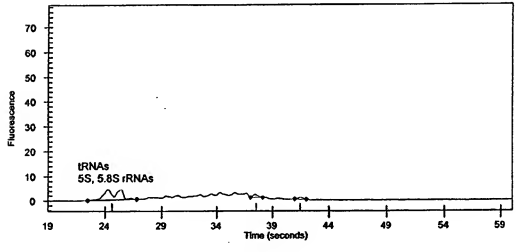

tRNAs
5S, 5.8S rRNAs
Fluorescence
Time (seconds)
70
60
50
40
30
20
10
0
19
24
29
34
39
44
49
54
59

FIG. 8B

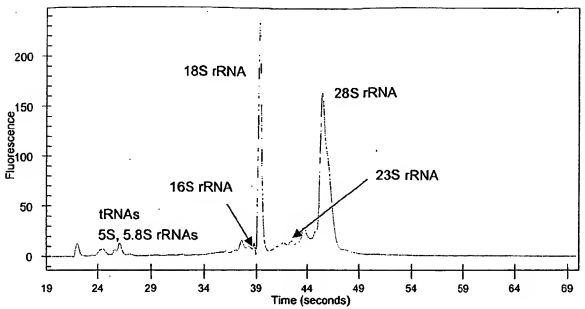

18S rRNA
28S rRNA
23S rRNA
16S rRNA
tRNAs
5S, 5.8S rRNAs
Fluorescence
Time (seconds)

FIG. 9A

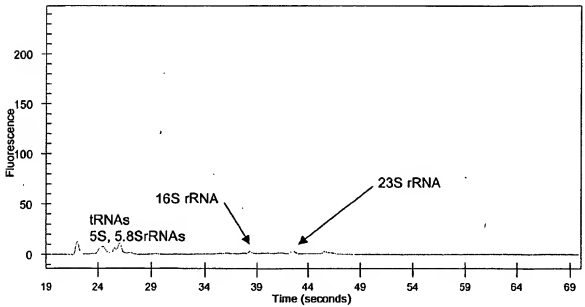

23S rRNA
16S rRNA
tRNAs
5S, 5.8SrRNAs
Fluorescence
Time (seconds)

FIG. 9B

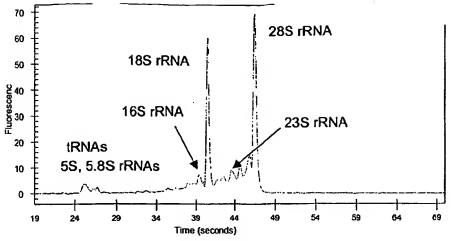


FIG. 10A

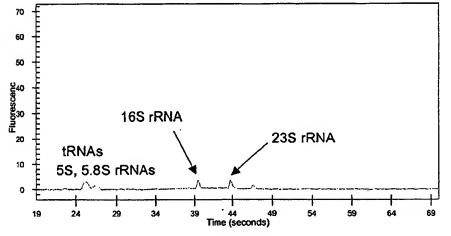


FIG. 10B

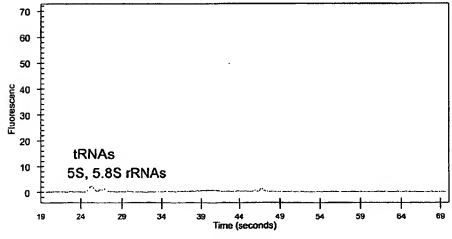


FIG. 10C

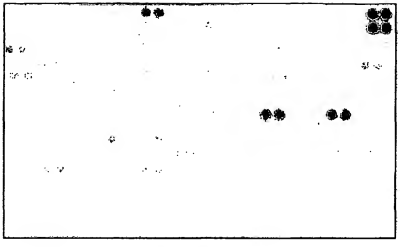


FIG. 11A

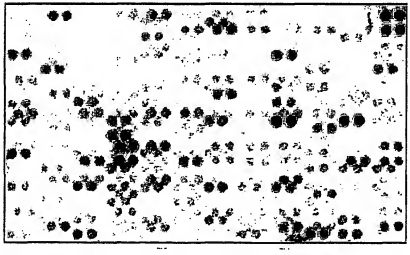


FIG. 11B

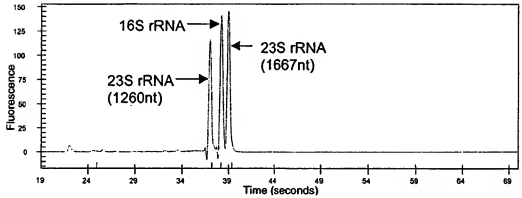


FIG. 12A

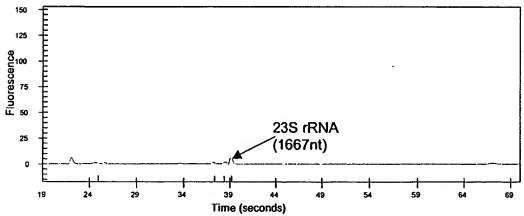


FIG. 12B

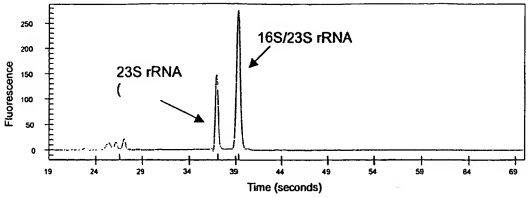


FIG. 13A

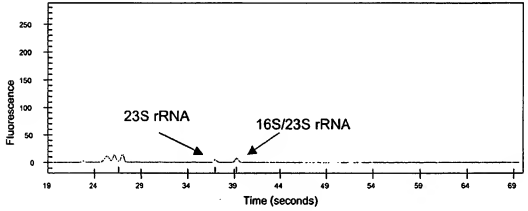


FIG. 13B

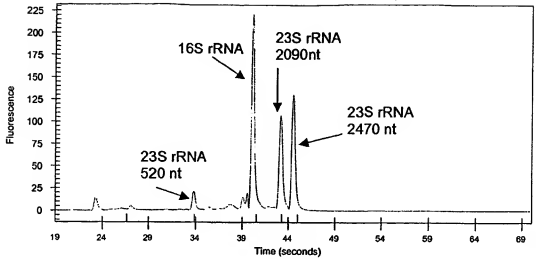


FIG 14A.

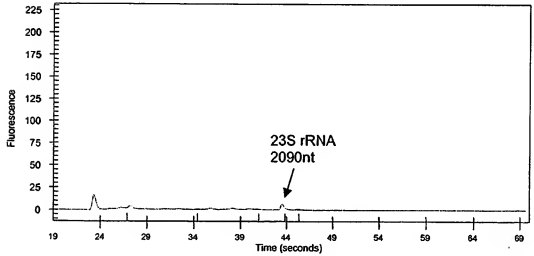


FIG 14B.

SEQUENCE LISTING

<110> MURPHY, GEORGE L.
WHITLEY, J. PENN

<120> METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS

<130> AMBI:076WO

<140> UNKNOWN
<141> 2002-12-20

<150> 10/029,397
<151> 2001-12-20

<160> 92

<170> PatentIn Ver. 2.1

<210> 1
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 1
ctgctgcctc ccgtaggagt ct 22

<210> 2
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 2
cgtattaccg cggctgctgg cac 23

<210> 3
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 3
cgcccagtaa ttccgattaa cgc 23

<210> 4
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 4
tggactacca gggtatctaa tcc 23

<210> 5
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 5
gggttgcgct cgttgcggga ctt 23

<210> 6
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 6
taaggaggtg atccaaccgc agg 23

<210> 7
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 7
ggttcttttt cactcccctc gcc 23

<210> 8
<211> 23

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 8
gacccattat acaaaaggta cgc 23

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 9
gccccgttac atcttccgcg cag 23

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 10
cgacaaggaa tttcgctacc tta 23

<210> 11
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 11
cttacccgac aaggaatttc gc 22

<210> 12
<211> 23
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic Primer

<400> 12
gagccgacat cgaggtgcca aac 23

<210> 13
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 13
ggttaagcct cacggttcat t 21

<210> 14
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 14
ggaagcgcac ggca 14

<210> 15
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 15
ccccttctcc cgaagttacg ggg 23

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 16

gtgagctatt acgctttctt t 21

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 17
taccggccgt gcgtacttag aca 23

<210> 18
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 18
tgccctccaa tggatcctcg tta 23

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 19
ctacggaaac cttgttacga ctt 23

<210> 20
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 20
gagcactggg cagaaatcac atc 23

<210> 21

<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 21
gtttcttttc ctccgctgac taa 23

<210> 22
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 22
tcctcagcca agcacataca cca 23

<210> 23
<211> 1427
<212> DNA
<213> Bacillus subtilis

<220>
<221> modified_base
<222> (554)..(873)
<223> N = A, C, G or T/U

<400> 23
gagagtttga tcctggctca ggacgaacgc tggcggcgtg cctaatacat gcaagtcgag 60
cggacagatg ggagcttgct ccctgatgtt agcggcggac gggtgagtaa cacgtgggta 120
acctgcctgt aagactggga taactccggg aaaccggggc taataccgga tggttgtttg 180
aaccgcatgg ttcaaacata aaaggtggct tcggctacca cttacagatg gacccgcggc 240
gcattagcta gttggtgagg taacggctca ccaaggcaac gatgcgtagc cgacctgaga 300
gggtgatcgg ccacactggg actgagacac ggcccagact cctacgggag gcagcagtag 360
ggaatcttcc gcaatggacg aaagtctgac ggagcaacgc cgcgtgagtg atgaaggttt 420
tcggatcgta aagctctgtt gttagggaag aacaagtacc gttcgaatag ggcggtacct 480
tgacggtacc taaccagaaa gccacggcta actacgtgcc agcagccgcg gtaatacgta 540
ggtggcaagc gttntccgga attattgggc gtaaagggct cgcaggcggt ttcttaagtc 600
tgatgtgaaa gccccggct caaccggga gggtcattgg aaactgggga acttgagtgc 660
agaagaggag agtggaattc cacgtgtngc ggtgaaatgc gtagagatgt ggaggaacac 720
cagtggcgaa ggcgactctc tggtctgtaa ctgacgctga ggagcgaaag cgtggggagc 780
gaacaggatt agataccctg gtagtccacg ccgtaaacga tgagtgctaa gtgttagggg 840
gtttccgccc cttagtgctg cagtaacgca ttnagcactc cgcctgggga gtacggtcgc 900
aagactgaaa ctcaaaggaa ttgacggggg ccgcacaagc ggtggagcat gtggtttaat 960
tcgaagcaac gcgaagaacc ttaccaggtc ttgacatcct ctgacaatcc tagagatagg 1020
acgtcttcgg gggcagagtg acaggtggtg catggttgtc gtcagctcgt gtcgtgagat 1080
gttgggttaa gtcccgcaac gagcgcaacc ctggatctta gttgccagca ttcagttggg 1140
cactctaagg tgactgccgg tgacaaaccg gaggaaggtg gggatgacgt caaatcatca 1200

tgccccttat gacctgggct acacacgtgc tacaatggac agaacaaagg gcagcgaaac 1260
cgcgaggtta agccaatccc acaaatctgt tctcagttcg gatcgcagtc tgcaactcga 1320
ctgcgtgaag ctggaatcgc tagtaatcgc ggatcagcat gccgcggtga atacgttccc 1380
gggccttgta cacaccgccc gtcacaccac gagagtttgt aacaccc 1427

<210> 24
<211> 1544
<212> DNA
<213> Bacillus anthracis

<400> 24
gtttgatcct ggctcaggat gaacgctggc ggcgtgccta atacatgcaa gtcgagcgaa 60
tggattaaga gcttgctctt atgaagttag cggcggacgg gtgagtaaca cgtgggtaac 120
ctgccctaaa gactgggata actccgggaa accgggctaa taccgataca acattttgaa 180
ccgcatggtt cgaaattgaa aggcggcttc ggctgtcact tatggatgga cccgcgtcgc 240
attagctagt tggtgaggta acggctcacc aaggcaacga tgcgtagccg acctgagagg 300
gtgatcggcc acactgggac tgagacacgg cccagactcc tacgggaggc agcagtaggg 360
aatcttccgc aatggacgaa agtctgacgg agcaacgccg cgtgagtgat gaaggcttcc 420
gggtcgtaaa actctgttgt tagggaagaa caagtgctag ttgaataagc tggcaccttg 480
acggtaccta accagaaagc cacggctaac tacgtgccag cagccgcggt aatacgtagg 540
tggcaagcgt tatccggaat tattgggcgt aaagcgcgcg caggtggttt cttaagtctg 600
atgtgaaagc ccacggctca accgtggagg gtcattggaa actgggagac ttgagtgcag 660
aagaggaaag tggaattcca tgtgtagcgg tgaaatgcgt agagatatgg aggaacacca 720
gtggcgaagg cgactttctg gtctgtaact gacactgagg cgcgaaagcg tggggagcaa 780
acaggattag ataccctggt agtccacgcc gtaaacgatg agtgctaagt gttagagggt 840
ttccgccctt tagtgctgaa gttaacgcat taagcactcc gcctggggag tacggccgca 900
aggctgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagcat gtggtttaat 960
tcgaagcaac gcgaagaacc ttaccaggtc ttgacatcct ctgacaaccc tagagatagg 1020
gcttctcctt cgggagcaga gtgacaggtg gtgcatggtt gtcgtcagct cgtgtcgtga 1080
gatgttgggt taagtcccgc aacgagcgca acccttgatc ttagttgcca tcattaagtt 1140
gggcactcta aggtgactgc cggtgacaaa ccggaggaag gtggggatga cgtcaaatca 1200
tcatgccccc tatgacctgg gctacacacg tgctacaatg gacggtacaa agagctgcaa 1260
gaccgcgagg tggagctaat ctcataaaac cgttctcagt tcggattgta ggctgcaact 1320
cgcctacatg aagctggaat cgctagtaat cgcggatcag catgccgcgg tgaatacgtt 1380
cccgggcctt gtacacaccg cccgtcacac cacgagagtt tgtaacaccc gaagtcggtg 1440
gggtaacctt tttggagcca gccgcctaag gtgggacaga tgattggggt gaagtcgtaa 1500
caaggtagcc gtatcggaag gtgcggctgg atcacctcct ttct 1544

<210> 25
<211> 1449
<212> DNA
<213> Enterococcus faecalis

<400> 25
cgaacgctgg cggcgtgcct aatacatgca agtcgaacgc ttctttcctc ccgagtgctt 60
gcactcaatt ggaaagagga gtggcggacg ggtgagtaac acgtgggtaa cctacccatc 120
agagggggat aacacttgga aacaggtgct aataccgcat aacagtttat gccgcatggc 180
ataagagtga aaggcgcttt cgggtgtcgc tgatggatgg acccgcggtg cattagctag 240
ttggtgaggt aacggctcac caaggccacg atgcatagcc gacctgagag ggtgatcggc 300
cacactggga ctgagacacg gcccagactc ctacgggagg cagcagtagg gaatcttcgg 360
caatggacga aagtctgacc gagcaacgcc gcgtgagtga agaaggtttt cggatcgtaa 420
aactctgttg ttagagaaga acaaggacgt tagtaactga acgtcccctg acggtatcta 480
accagaaagc cacggctaac tacgtgccag cagccgcggt aatacgtagg tggcaagcgt 540
tgtccggatt tattgggcgt aaagcgagcg caggcggttt cttaagtctg atgtgaaagc 600

ccccgctca accggggagg gtcattggaa actgggagac ttgagtgcag aagaggagag 660
tggaattcca tgtgtagcgg tgaaatgcgt agatatatgg aggaacacca gtggcgaagg 720
cggctttctg gtctgtaact gacgctgagg ctcgaaagcg tggggagcaa acaggattag 780
ataccctggt agtccacgcc gtaaacgatg agtgctaagt gttggagggt ttccgcctt 840
cagtgctgca gcaaacgcat taagcactcc gcctggggag tacgaccgca aggttgaaac 900
tcaaaggaat tgacgggggc ccgcacaagc ggtggagcat gtggtttaat tcgaagcaac 960
gcgaagaacc ttaccaggtc ttgacatcct ttgaccactc tagagataga gctttccctt 1020
cggggacaaa gtgacaggtg gtgcatggtt gtcgtcagct cgtgtcgtga gatgttgggt 1080
taagtcccgc aacgagcgca acccttattg ttagttgcca tcatttagtt gggcactcta 1140
gcgagactgc cggtgacaaa ccggaggaag gtggggatga cgtcaaatca tcatgcccct 1200
tatgacctgg gctacacacg tgctacaatg ggaagtacaa cgagtcgcta gaccgcgagg 1260
tcatgcaaat ctcttaaagc ttctctcagt tcggattgca ggctgcaact cgcctgcatg 1320
aagccggaat cgctagtaat cgcggatcag cacgccgcgg tgaatacgtt cccgggcctt 1380
gtacacaccg cccgtcacac cacgagagtt tgtaacaccc gaagtcggtg aggtaacctt 1440
tttggagcc 1449

<210> 26
<211> 1548
<212> DNA
<213> Lactococcus lactis

<400> 26
tttatttgag agtttgatcc tggctcagga cgaacgctgg cggcgtgcct aatacatgca 60
agttgagcgc tgaaggttgg tacttgtacc gactggatga gcagcgaacg ggtgagtaac 120
gcgtggggaa tctgcctttg agcgggggac aacatttgga aacgaatgct aataccgcat 180
aaaaacttta aacacaagtt ttaagtttga aagatgcaat tgcatcactc aaagatgatc 240
ccgcgttgta ttagctagtt ggtgaggtaa aggctcacca aggcgatgat acatagccga 300
cctgagaggg tgatcggcca cattgggact gagacacggc ccaaactcct acgggaggca 360
gcagtaggga atcttcggca atggacgaaa gtctgaccga gcaacgccgc gtgagtgaag 420
aaggttttcg gatcgtaaaa ctctgttggt agagaagaac gttggtgaga gtggaaagct 480
catcaagtga cggtaactac ccagaaaggg acggctaact acgtgccagc agccgcggta 540
atacgtaggt cccgagcgtt gtccggattt attgggcgta aagcgagcgc aggtggttta 600
ttaagtctgg tgtaaaaggc agtggctcaa ccattgtatg cattggaaac tggtagactt 660
gagtgcagga gaggagagtg gaattccatg tgtagcggtg aaatgcgtag atatatggag 720
gaacaccggt ggcgaaagcg gctctctggc ctgtaactga cactgaggct cgaaagcgtg 780
gggagcaaac aggattagat accctggtag tccacgccgt aaacgatgag tgctagatgt 840
agggagctat aagttctctg tatcgcagct aacgcaataa gcactccgcc tggggagtac 900
gaccgcaagg ttgaaactca aaggaattga cgggggcccg cacaagcggt ggagcatgtg 960
gtttaattcg aagcaacgcg aagaacctta ccaggtcttg acatactcgt gctattccta 1020
gagataggaa gttccttcgg gacacgggat acaggtggtg catggttgtc gtcagctcgt 1080
gtcgtgagat gttgggttaa gtcccgcaac gagcgcaacc cctattgtta gttgccatca 1140
ttaagttggg cactctaacg agactgccgg tgataaaccg gaggaaggtg gggatgacgt 1200
caaatcatca tgccccttat gacctgggct acacacgtgc tacaatggat ggtacaacga 1260
gtcgcgagac agtgatgttt agctaatctc ttaaaaccat tctcagttcg gattgtaggc 1320
tgcaactcgc ctacatgaag tcggaatcgc tagtaatcgc ggatcagcac gccgcggtga 1380
atacgttccc gggccttgta cacaccgccc gtcacaccac gggagttggg agtacccgaa 1440
gtaggttgcc taaccgcaag gagggcgctt cctaaggtaa gaccgatgac tggggtgaag 1500
tcgtaacaag gtagccgtat cggaaggtgc ggctggatca cctccttt 1548

<210> 27
<211> 1524
<212> DNA
<213> Listeria monocytogenes

<400> 27
gcctgcaggt cgacaacaga gtttgatcat ggctcaggac gaacgctggc ggcgtgccta 60
atacatgcaa gtcgaacgaa cggaggaaga gcttgctctt ccaaagttag tggcggacgg 120
gtgagtaaca cgtgggcaac ctgcctgtaa gttgggggata actccgggaa accgggctca 180
ataccgaatg ataaagtgtg gcgcatgcca cgctttgaa agatggtttc ggctatcgct 240
tacagatggg cccgcggtgc attagctagt tggtagggta atggcctacc aaggcaacga 300
tgcatagccg acctgagagg gtgatcggcc acactgggac tgagacacgg cccagactcc 360
tacgggaggc agcagtaggg aatcttccgc aatggacgaa agtctgacgg agcaacgccg 420
cgtgtatgaa gaaggtttcc ggatcgtaaa gtactgttgt tagagaagaa caaggataag 480
agtaactgct tgtcccttga cggtatctaa ccagaaagcc acggctaact acgtgccagc 540
agccgcggta atacgtaggt ggcaagcgtt gtccgatttt attgggcgta aagcgcgcgc 600
aggcggtctt ttaagtctga tgtgaaagcc cccggcttaa ccggggaggg tcattggaaa 660
ctggaagact ggagtgcaga agaggagagt ggaattccac gtgtagcggt gaaatgcgta 720
gatatgtgga ggaacaccag tggcgaaggc gactctctgg tctgtaactg acgctgaggc 780
gcgaaagcgt ggggagcaaa caggattaga taccctggta gtccacgccg taaacgatga 840
gtgctaagtg ttaggggtt tccgcccctt agtgctgcag ctaacgcatt aagcactctg 900
cctggggagt acgaccgcaa ggttgaaact caaaggaatt gacgggggcc cgcacaagcg 960
tggagcatgt ggtttaattc gaagcaacgc gaagaacctt accaggtctt gacatccttt 1020
gaccactctg gagacagagc tttccttcg ggacaaagtg acaggtggtg catggttgtc 1080
gtcagctcgt gtcgtgagat gttgggttaa gtcccgcaac gagcgcaacc cttgatttta 1140
gttgccagca tttagttggg cactctaaag tgactgccgg tgcaaggcga ggaaggtggg 1200
gatgacgtca aatcatcatg ccccttatga cctgggctac acacgtgcta caatggatag 1260
tacaaagggt cgcgaagccg cgaggtggag ctaatcccat aaaactattc tcagttcgga 1320
ttgtaggctg caactcgcct acatgaagcc ggaatcgcta gtaatcgtgg atcagcatgc 1380
cacggtgagt acgttcccgg gccttgtaca caccgcccgt cacacccaga gagtttgtaa 1440
cacccgaagt cggtaggga acctttatgg agccagccgc cgaaggtggg acagataatt 1500
ggggtgaagt cgtaacaagg taaa 1524

<210> 28
<211> 1555
<212> DNA
<213> Staphylococcus aureus

<400> 28
ttttatggag agtttgatcc tggctcagga tgaacgctgg cggcgtgcct aatacatgca 60
agtcgagcga acggacgaga agcttgcttc tctgatgtta gcggcggacg ggtgagtaac 120
acgtggataa cctacctata agactgggat aacttcggga aaccggagct aataccggat 180
aatattttga accgcatggt tcaaaagtga aagacggtct tgctgtcact atagatggga 240
tccgcgctgc attagctagt tggtaaggta acggcttacc aaggcaacga tacgtagccg 300
acctgagagg gtgatcggcc acactggaac tgagacacgg tccagactcc tacgggaggc 360
agcagtaggg aatcttccgc aatgggcgaa agcctgacgg agcaacgccg cgtgagtgat 420
gaaggttcttc ggatcgtaaa actctgttat tagggaagaa catatgtgta agtaactgtg 480
cacatcttga cggtacctaa tcagaaagcc acggctaact acgtgccagc agccgcggta 540
atacgtaggt ggcaagcgtt atccggaatt attgggcgta aagcgcgcgt aggcggtttt 600
ttaagtctga tgtgaaagcc cacggctcaa ccgtggaggg tcattggaaa ctggaaaact 660
tgagtgcaga agaggaaagt ggaattccat gtgtagcggt gaaatgcgca gagatatgga 720
ggaacaccag tggcgaaggc gactttctgg tctgtaactg acgctgatgt gcgaaagcgt 780
ggggatcaaa caggattaga taccctggta gtccacgccg taaacgatga gtgctaagtg 840
ttaggggtt tccgcccctt agtgctgcag ctaacgcatt aagcactccg cctggggagt 900
acgaccgcaa ggttgaaact caaaggaatt gacggggacc cgcacaagcg gtggagcatg 960
tggtttaatt cgaagcaacg cgaagaacct taccaaatct tgacatcctt tgacaactct 1020
agagatagag ccttcccctt cgggggacaa agtgacaggt ggtgcatggt tgtcgtcagc 1080
tcgtgtcgtg agatgttggg ttaagtcccg caacgagcgc aaccctaag cttagttgcc 1140
atcattaagt tgggcactct aagttgactg ccggtgacaa accggaggaa ggtggggatg 1200
acgtcaaatc atcatgcccc ttatgatttg ggctacacac gtgctacaat ggacaataca 1260

```
aagggcagcg aaaccgcgag gtcaagcaaa tcccataaag ttgttctcag ttcggattgt 1320
agtctgcaac tcgactacat gaagctggaa tcgctagtaa tcgtagatca gcatgctacg 1380
gtgaatacgt tcccgggtat tgtacacacc gcccgtcaca ccacgagagt ttgtaacacc 1440
cgaagccggt ggagtaacct tttaggagct agccgtcgaa ggtgggacaa atgattgggg 1500
tgaagtcgta acaaggtagc cgtatcggaa ggtgcggctg gatcacctcc tttct      1555
```

```
<210> 29
<211> 1551
<212> DNA
<213> Streptococcus mutans

<400> 29
agagtttgat cctggctcag gacgaacgct ggcggcgtgc ctaatacatg caagtgggac 60
gcaaggaaac acactgtgct tgcacaccgt gttttcttga gtcgcgaacg ggtgagtaac 120
gcgtaggtaa cctgcctatt agcgggggat aactattgga aacgatagct aatacccgcat 180
aatattaatt attgcatgat aattgattga aagatgcaag cgcatcacta gtagatggac 240
ctgcgttgta ttagctagtt ggtaaggtaa gagcttacca aggcgacgat acatagccga 300
cctgagaggg tgatcggcca cactgggact gagacacggc ccagactcct acgggaggca 360
gcagtaggga atcttcggca atggacgaaa gtctgaccga gcaacgccgc gtgagtgaag 420
aaggttttcg gatcgtaaag ctctgttgta agtcaagaac gtgtgtgaga gtggaaagtt 480
cacacagtga cggtagctta ccagaaaggg acggctaact acgtgccagc agccgcggta 540
atacgtaggt cccgagcgtt gtccggattt attgggcgta aagggagcgc aggcggtcag 600
gaaagtctgg agtaaaaggc tatggctcaa ccatagtgtg ctctggaaac tgtctgactt 660
gagtgcagaa ggggagagtg gaattccatg tgtagcggtg aaatgcgtag atatatggag 720
gaacaccagt ggcgaaagcg gctctctggt ctgtcactga cgctgaggct cgaaagcgtg 780
ggtagcgaac aggattagat accctggtag tccacgccgt aaacgatgag tgctaggtgt 840
taggcccttt ccggggctta gtgccggagc taacgcaata agcactccgc ctggggagta 900
cgaccgcaag gttgaaactc aaaggaattg acgggggccc gcacaagcgg tggagcatgt 960
ggtttaattc gaagcaacgc gaagaacctt accaggtctt gacatcccga tgctattctt 1020
agagatagga agttacttcg gtacatcgga gacaggtggt gcatggttgt cgtcagctcg 1080
tgtcgtgaga tgttgggtta agtcccgcaa cgagcgcaac ccttattgtt agttgccatc 1140
attaagttgg gcactctagc gagactgccg gtaataaacc ggaggaaggt ggggatgacg 1200
tcaaatcatc atgcccctta tgacctgggc tacacacgtg ctacaatggt cggtacaacg 1260
agttgcgagc cggtgacggc aagctaatct ctgaaagccg atctcagttc ggattggagg 1320
ctgcaactcg cctccatgaa gtcggaatcg ctagtaatcg cggatcagca cgccgcggtg 1380
aatacgttcc cgggccttgt acacaccgcc cgtcacacca cgagagtttg taacacccga 1440
agtcggtgag gtaacctttt aaggggcaag ccgcctaagg tgggatggat gattggggtg 1500
aagtcgtaac aaggtagccg tatcggaagg tgcggctgga tcacctcctt t           1551
```

```
<210> 30
<211> 1515
<212> DNA
<213> Streptococcus pneumoniae

<400> 30
atttgatcct ggctcaggac gaacgctggc ggcgtgccta atacatgcaa gtagaacgct 60
gaaggaggag cttgcttctc tggatgagtt gcgaacgggt gagtaacgcg taggtaacct 120
gcctggtagc gggggataac tattggaaac gatagctaat accgcataag agtggatgtt 180
gcatgacatt tgcttaaaag gtgcacttgc atcactacca gatggacctg cgttgtatta 240
gctagttggt ggggtaacgg ctcaccaagg cgacgataca tagccgacct gagagggtga 300
tcggccacac tgggactgag acacgkccca gactcctacg ggaggcagca gtagggaatc 360
ttcggcaatg gacggaagtc tgaccgagca acgccgcgtg agtgaagaag gttttcggat 420
cgtaaagctc tgttgtaaga gaagaacgag tgtgagagtg gaaagttcac actgtgacgg 480
tatcttacca gaaagggacg gctaactacg tgccagcagc cgcggtaata cgtaggtccc 540
```

gagcgttgtc cggatttatt gggcgtaaag cgagcgcagg cggttagata agtctgaagt 600
taaaggctgt ggcttaacca tagtaggctt tggaaactgt ttaacttgag tgcaagaggg 660
gagagtggaa ttccatgtgt agcggtgaaa tgcgtagata tatggaggaa caccggtggc 720
gaaaggcggct ctctggcttg taactgacgc tgaggctcga aagcgtgggg agcaaacagg 780
attagatacc ctggtagtcc acgctgtaaa cgatgagtgc taggtgttag accctttccg 840
gggtttagtg ccgtagctaa cgcattaagc actccgcctg gggagtacga ccgcaaggtt 900
gaaactcaaa ggaattgacg gggcccgca caagcggtgg agcatgtggt ttaattcgaa 960
gcaacgcgaa gaaccttacc aggtcttgac atcctctcga ccgctctaga gatagagttt 1020
tccttcggga cagaggtgac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt 1080
tgggttaagt cccgcaacga gcgcaacccc tattgttagt tgccatcatt cagttgggca 1140
ctctagcgag actgccggta ataaaccgga ggaaggtggg gatgacgtca aatcatcatg 1200
ccccttatga cctgggctac acacgtgcta caatggctgg tacaacgagt cgcaagccgg 1260
tgacggcaag ctaatctctt aaagccagtc tcagttcgga ttgtaggctg caactcgcct 1320
acatgaagtc ggaatcgcta gtaatcgcgg atcagcacgc cgcggtgaat acgttcccgg 1380
gccttgtaca caccgcccgt cacaccacga gagtttgtaa cacccgaagt cggtgaggta 1440
accgtaagga gccagccgcc taaggtggga tagatgattg gggtgaagtc gtaacaaggt 1500
cagccgtttg ggaga 1515

<210> 31
<211> 1335
<212> DNA
<213> Streptococcus pyogenes

<400> 31
gaacgggtga gtaacgcgta ggtaacctac ctcatagcgg gggataacta ttggaaacga 60
tagctaatac cgcataagag agactaacgc atgttagtaa tttaaaaggg gcaattgctc 120
cactatgaga tggacctgcg ttgtattagc tagttggtga ggtaaaggct caccaaggcg 180
acgatacata gccgacctga gagggtgatc ggccacactg ggactgagac acggcccaga 240
ctcctacggg aggcagcagt agggaatctt cggcaatggg ggcaaccctg accgagcaac 300
gccgcgtgag tgaagaaggt tttcggatcg taaagctctg ttgttagaga agaatgatgg 360
tgggagtgga aaatccacca agtgacggta actaaccaga aagggacggc taactacgtg 420
ccagcagccg cggtaatacg taggtcccga gcgttgtccg gatttattgg gcgtaaagcg 480
agcgcaggcg gtttttaag tctgaagtta aaggcattgg ctcaaccaat gtacgctttg 540
gaaactggaa aacttgagtg cagaagggga gagtggaatt ccatgtgtag cggtgaaatg 600
cgtagatata tggaggaaca ccggtggcga aagcggctct ctggtctgta actgacgctg 660
aggctcgaaa gcgtggggag caaacaggat tagataccct ggtagtccac gccgtaaacg 720
atgagtgcta ggtgttaggc cctttccggg gcttagtgcc ggagctaacg cattaagcac 780
tccgcctggg gagtacgacc gcaaggttga aactcaaagg aattgacggg ggcccgcaca 840
agcggtggag catgtggttt aattcgaagc aacgcgaaga accttaccag gtcttgacat 900
cccgatgccc gctctagaga tagagtttta cttcggtaca tcggtgacag gtggtgcatg 960
gttgtcgtca gctcgtgtcg tgagatgttg ggttaagtcc cgcaacgagc gcaaccccta 1020
ttgttagttg ccatcattaa gttgggcact ctagcgagac tgccggtaat aaaccggagg 1080
aaggtgggga tgacgtcaaa tcatcatgcc ccttatgacc tgggctacac acgtgctaca 1140
atggttggta caacgagtcg caagccggtg acggcaagct aatctcttaa agccaatctc 1200
agttcggatt gtaggctgca actcgcctac atgaagtcgg aatcgctagt aatcgcggat 1260
cagcacgccg cggtgaatac gttcccgggc cttgtacaca ccgcccgtca caccacgaga 1320
gtttgtaaca cccga 1335

<210> 32
<211> 1465
<212> DNA
<213> Mycobacterium avium

<220>

<221> modified_base
<222> (298)..(881)
<223> N = A, C, G or T/U

<400> 32
ggcggcgtgc ttaacacatg caagtcgaac ggaaaggcct cttcggaggt actcgagtgg 60
cgaacgggtg agtaacacgt gggcaatcta ccctgcactt cgggataagc ctgggaaact 120
gggtctaata ccggatagga cctcaagacg catgtcttct ggtggaaagc ttttgcggtg 180
tggggatggg ccgcggccta tcagcttgtt ggtggggtga cggcctacca aggcgacgac 240
gggtagccgg cctgagaggg tgtccggcca cactgggact gagatacggc ccagactnct 300
acgggaggca gcagtgggga atattgcaca atgggcgcaa gcctgatgca gcgacgccgc 360
gtgggggatg acggccttcg ggttgtaaac ctctttcacc atcgacgaag gtccgggttt 420
tctcggattg acggtaggtg gagaagaagc accggccaac tacgtgccag cagccgcggt 480
aatacgtagg gtgcgagcgt tgtccggaat tactgggcgt aaagagctcg taggtggttt 540
gtcgcgttgt tcgtgaaatc tcacggctta actgtgagcg tgcgngcgat acgggcagac 600
tagagtactg caggggagac tggaattcct ggtgtagcgg tggaatgcgc agatatcagg 660
aggaacaccg gtggcgaagg cgggtctctg ggcagtaact gacgctgagg agcgaaagcg 720
tggggagcga acaggattag ataccctggt agtccacgnc gtaaacggtg ggtactaggt 780
gtgggtttcc ttccttggga tccgtgccgt agctaacgca ttaagtaccc cgcctgggga 840
gtacggncgc aaggctaaaa ctcaaaggaa ttgacggggg nccgcacaag cggcggagca 900
tgtggattaa ttcgatgcaa cgcgaagaac cttacctggg tttgacatgc acaggacgcg 960
tctagagata ggcgttccct tgtggcctgt gtgcaggtgg tgcatggctg tcgtcagctc 1020
gtgtcgtgag atgttgggtt aagtcccgca acgagcgcaa cccttgtctc atgttgccag 1080
cgggtaatgc cggggactcg tgagagactg ccggggtcaa ctcggaggaa ggtggggatg 1140
acgtcaagtc atcatgcccc ttatgtccag ggcttcacac atgctacaat ggccggtaca 1200
aagggctgcg atgccgtaag gttaagcgaa tccttttaaa gccggtctca gttcggattg 1260
gggtctgcaa ctcgacccca tgaagtcgga gtcgctagta atcgcagatc agcaacgctg 1320
cggtgaatac gttcccgggc cttgtacaca ccgcccgtca cgtcatgaaa gtcggtaaca 1380
cccgaagcca gtggcctaac ccttttggga gggagctgtc gaaggtggga tcggcgattg 1440
ggacgaagtc gtaacaaggt agccg 1465

<210> 33
<211> 1536
<212> DNA
<213> Mycobacterium tuberculosis

<400> 33
tttgtttgga gagtttgatc ctggctcagg acgaacgctg gcggcgtgct taacacatgc 60
aagtcgaacg gaaaggtctc ttcggagata ctcgagtggc gaacgggtga gtaacacgtg 120
ggtgatctgc cctgcacttc gggataagcc tgggaaactg ggtctaatac cggatggcac 180
cacgggatgc atgtcttgtg gtggaaagcg ctttagcggt gtgggatgag cccgcggcct 240
atcagcttgt tggtggggtg acggcctacc aaggcgacga cgggtagccg gcctgagagg 300
gtgtccggcc acactgggac tgagatacgg cccagactcc tacgggaggc agcagtgggg 360
aatattgcac aatgggcgca agcctgatgc agcgacgccg cgtgggggat gacggccttc 420
gggttgtaaa cctctttcac catcgacgaa ggtccgggtt ctctcggatt gacggtaggt 480
ggagaagaag caccggccaa ctacgtgcca gcagccgcgg taatacgtag ggtgcgagcg 540
ttgtccggaa ttactgggcg taaagagctc gtaggtggtt tgtcgcgttg ttcgtgaaat 600
ctcacggctt aactgtgagc gtgcgggcga tacgggcaga ctagagtact gcagggaga 660
ctggaattcc tggtgtagcg gtggaatgcg cagatatcag gaggaacacc ggtggcgaag 720
gcgggtctct gggcagtaac tgacgctgag gagcgaaagc gtggggagcg aacaggatta 780
gataccctgg tagtccacgc cgtaaacggt gggtactagg tgtgggtttc cttccttggg 840
atccgtgccg tagctaacgc attaagtacc ccgcctgggg agtacggccg caaggctaaa 900
actcaaagga attgacgggg gcccgcacaa gcggcggagc atgtggatta attcgatgca 960
acgcgaagaa ccttacctgg gtttgacatg cacaggacgc gtctagagat aggcgttccc 1020
ttgtggcctg tgtgcaggtg gtgcatggct gtcgtcagct cgtgtcgtga gatgttgggt 1080

taagtcccgc aacgagcgca acccttgtct catgttgcca gcacgtaatg gtggggactc 1140
gtgagagact gccggggtca actcggagga aggtggggat gacgtcaagt catcatgccc 1200
cttatgtcca gggcttcaca catgctacaa tggccggtac aaagggctgc gatgccgcga 1260
ggttaagcga atccttaaaa gccggtctca gttcggatcg gggtctgcaa ctcgaccccg 1320
tgaagtcgga gtcgctagta atcgcagatc agcaacgctg cggtgaatac gttcccgggc 1380
cttgtacaca ccgcccgtca cgtcatgaaa gtcggtaaca cccgaagcca gtggcctaac 1440
cctcgggagg gagctgtcga aggtgggatc ggcgattggg acgaagtcgt aacaaggtag 1500
ccgtaccgga aggtgcggct ggatcacctc ctttct 1536

<210> 34
<211> 1536
<212> DNA
<213> Escherichia coli

<400> 34
tttgtttgga gagtttgatc ctggctcagg acgaacgctg gcggcgtgct taacacatgc 60
aagtcgaacg gaaaggtctc ttcggagata ctcgagtggc gaacgggtga gtaacacgtg 120
ggtgatctgc cctgcacttc gggataagcc tgggaaactg ggtctaatac cggataggac 180
cacgggatgc atgtcttgtg gtggaaagcg ctttagcggt gtgggatgag cccgcggcct 240
atcagcttgt tggtggggtg acggcctacc aaggcgacga cgggtagccg gcctgagagg 300
gtgtccggcc acactgggac tgagatacgg cccagactcc tacgggaggc agcagtgggg 360
aatattgcac aatgggcgca agcctgatgc agccatgccg cgtggggat gacggccttc 420
gggttgtaaa cctctttcac catcgacgaa ggtccgggtt ctctcggatt gacggtaggt 480
ggagaagaag caccggccaa ctacgtgcca gcagccgcgg taatacgtag ggtgcgagcg 540
ttgtccggaa ttactgggcg taaagagctc gtaggtggtt tgtcgcgttg ttcgtgaaat 600
ctcacggctt aactgtgagc gtgcgggcga tacgggcaga ctagagtact gcaggggaga 660
ctggaattcc tggtgtagcg gtggaatgcg cagatatcag gaggaacacc ggtggcgaag 720
gcgggtctct gggcagtaac tgacgctgag gagcgaaagc gtggggagcg aacaggatta 780
gataccctgg tagtccacgc cgtaaacggt gggtactagg tgtgggtttc cttccttggg 840
atccgtgccg tagctaacgc attaagtacc ccgcctgggg agtacggccg caaggctaaa 900
actcaaagga attgacgggg gcccgcacaa gcggcggagc atgtggatta attcgatgca 960
acgcgaagaa ccttacctgg gtttgacatg cacaggacgc gtctagagat aggcgttccc 1020
ttgtggcctg tgtgcaggtg gtgcatggct gtcgtcagct cgtgtcgtga gatgttgggt 1080
taagtcccgc aacgagcgca acccttgtct catgttgcca gcacgtaatg gtggggactc 1140
gtgagagact gccggggtca actcggagga aggtggggat gacgtcaagt catcatgccc 1200
cttatgtcca gggcttcaca catgctacaa tggccggtac aaagggctgc gatgccgcga 1260
ggttaagcga atccttaaaa gccggtctca gttcggatcg gggtctgcaa ctcgaccccg 1320
tgaagtcgga gtcgctagta atcgcagatc agcaacgctg cggtgaatac gttcccgggc 1380
cttgtacaca ccgcccgtca cgtcatgaaa gtcggtaaca cccgaagcca gtggcctaac 1440
cctcgggagg gagctgtcga aggtgggatc ggcgattggg acgaagtcgt aacaaggtag 1500
ccgtaccgga aggtgcggct ggatcacctc ctttct 1536

<210> 35
<211> 1534
<212> DNA
<213> Klebsiella pneumoniae

<220>
<221> modified_base
<222> (11)..(12)
<223> N = A, C, G or T/U

<400> 35
agagtttgat nntggctcag attgaacgct ggcggcaggc ctaacacatg caagtcgagc 60

ggtagcacag agagcttgct ctcgggtgac gagcggcgga cgggtgagta atgtctggga 120
aactgcctga tggaggggga taactactgg aaacggtagc taataccgca taacgtcgca 180
agaccaaagt gggggacctt cgggcctcat gccatcagat gtgcccagat gggattagct 240
agtaggtggg gtaacggctc acctaggcga cgatccctag ctggtctgag aggatgacca 300
gccacactgg aactgagaca cggtccagac tcctacggga ggcagcagtg gggaatattg 360
cacaatgggc gcaagcctga tgcagccatg ccgcgtgtgt gaagaaggcc ttcgggttgt 420
aaagcacttt cagcggggag gaaggcgatg aggttaataa cctcatcgat tgacgttacc 480
ctgcagaaga agcaccggct aactccgtgc cagcagccgc ggtaatacgg agggtgcaag 540
cgttaatcgg aattactggg cgtaaagcgc acgcaggcgg tctgtcaagt cggatgtgaa 600
atccccgggc tcaacctggg aactgcattc gaaactggca ggctagagtc ttgtagaggg 660
gggtagaatt ccaggtgtag cggtgaaatg cgtagagatc tggaggaata ccggtggcga 720
aggcggcccc ctggacaaag actgacgctc aggtgcgaaa gcgtggggag caaacaggat 780
tagataccct ggtagtccac gccgtaaacg atgtcgattt ggaggttgtg cccttgaggc 840
gtggcttccg gagctaacgc gttaaatcga ccgcctgggg agtacggccg caaggttaaa 900
actcaaatga attgacgggg gcccgcacaa gcggtggagc atgtggttta attcgatgca 960
acgcgaagaa ccttacctgg tcttgacatc cacagaactt tccagagatg gattggtgcc 1020
ttcgggaact gtgagacagg tgctgcatgg ctgtcgtcag ctcgtgttgt gaaatgttgg 1080
gttaagtccc gcaacgagcg caacccttat cctttgttgc cagcggttag gccgggaact 1140
caaaggagac tgccagtgat aaactggagg aaggtgggga tgacgtcaag tcatcatggc 1200
ccttacgacc agggctacac acgtgctaca atggcatata caaagagaag cgacctcgcg 1260
agagcaagcg gacctcataa agtatgtcgt agtccggatt ggagtctgca actcgactcc 1320
atgaagtcgg aatcgctagt aatcgtagat cagaatgcta cggtgaatac gttcccgggc 1380
cttgtacaca ccgcccgtca caccatggga gtgggttgca aaagaagtag gtagcttaac 1440
cttcgggagg gcgcttacca ctttgtgatt catgactggg gtgaagtcgt aacaaggtaa 1500
ccgtagggga acctgcggtt ggatcacctc cttt 1534

<210> 36
<211> 1485
<212> DNA
<213> ACTINOBACCILUS ACTIN

<220>
<221> modified_base
<222> (208)..(1476)
<223> N = A, C, G or T/U

<400> 36
attgaagagt ttgatcatgg ctcagattga acgctggcgg caggcttaac acatgcaagt 60
cggacggtag caggagaaag cttgctttct tgctgacgag tggcggacgg gtgagtaatg 120
cttgggaatc tgtcttatgg agggggataa cgacgggaaa ctgtcgctaa taccgcgtag 180
agtcgggaga cgaaagtgcg ggactttntg gccgcatgcc atgagatgag cccaagtgtg 240
attaggtagt tggtggggta aaggcctacc aagccgacga tcgctagctg gtctgagagg 300
atggccagcc acaccgggac tgagacacgg cccngactcc tacgggaggc agcagtgggg 360
aatattgcgc aatgggggca accctgacgc agccatgccg cgtgaatgaa gaaggccttc 420
gggttgtaaa gttctttcgg tattgaggaa ggttggtgtg ttaatagcat gccaaattga 480
cgttaaatac agaagaagca ccggctaact ccgtgccagc agccgcggta atacgggggg 540
tgcgagcgtt aatcggaata actgggcgta aagggcacgt aggcggacct ttaagtgagg 600
tgtgaaatcc ccgggcttaa cctgggnatt gcatttcata ctgggggtct ggagtacttt 660
ngggagggnt agaattccac gtgtagcggt gaaatgcgta gagatgtgga ggaataccga 720
aggcgaaggc agccccttgg ggatgtactg acgctgatgt gcgaaagcgt ggggagcaaa 780
caggattaga taccctggta gtccacgctg taaacggtgt cgatttgggg attgggggtt 840
agccctggtg cccgaagcta acgtgataaa tcgaccgcct ggggagtacg gccgcaaggt 900
taaaactcaa atgaattgac gggggcccgc acaagcggtg gagcatgtgg tttaattcga 960
tgcaacgcga agaaccttac ctactcttga catccgaaga agaactcaga gatgggttg 1020
tgccttaggg agctttgaga caggtgctgc atggcngtcg tcagctcgtg ttgtgaaatg 1080

ttgggttaag tcccgcaacg agcgcaaccc ttatcctttg tggccagcga cgtggtcggg 1140
aactcaaagg agactgccgg tgataaaccg gaggaaggtg gggatgacgt caagtcatca 1200
tggcccttac gagtagggct acacacgtgc tacaatggcg tatacagagg gtaacccaac 1260
agcgatgggg agtgaatctc agaaagtgcg tctaagttcg gattggagtc tgcaactcga 1320
ctccatgaag tcggaatcgc tagtaatcgc gaatcagaat gttgcggtga atacgttccc 1380
gggccttgta cacaccgccc gtcacaccat gggagtgggt tgtaccagaa gtggatagct 1440
gaaccgagag ggtggcgttt accacggtat gattcangac tgggg 1485

<210> 37
<211> 1487
<212> DNA
<213> Haemophilus influenzae

<220>
<221> modified_base
<222> (1)..(1387)
<223> N = A, C, G or T/U

<400> 37
naattgaaga gtttgatcat ggctcagatt gaacgctggc ggcaggctta acacatgcaa 60
gtcgaacggt agcaggagaa agcttgcttt cttgctgacg agtggcggac gggtgagtaa 120
tgcttgggaa tctggcttat ggagggggat aacgacggga aactgtcgct aataccgcgt 180
attatcggaa gatgaaagtg cgggactgag aggccgcatg ccataggatg agcccaagtg 240
ggattaggta gttggtgggg taaatgccta ccaagcctgc gatctctagc tggtctgaga 300
ggatgaccag ccacactgga actgagacac ggtccagact cctacgggag gcagcagtgg 360
ggaatattgc gcnatggggg gaaccctgac gcagccatgc cgcgtgaatg aagaaggcct 420
tcgggttgta aagttctttc ggtattgagg aaggttgatg tgttaatagc acatcaaatt 480
gacgttaaat acagaagaag caccggctaa ctccgtgcca gcagccgcgg taatacggag 540
ngtgcgagcg ttaatcggaa taactgggcg taaagggcac gcaggcggtt atttaagtga 600
ggtgtgaaag ccccgggctt aacctgggna ttgcatttca gactgggtaa ctagagtact 660
ttagggaggg gtagaattcc acgtgtagcg gtgaaatgcg tagagatgtg gaggaatacc 720
gaaggcgaag gcagcccctt gggaatgtac tgacgctcat gtgcgaaagc gtggggagca 780
aacaggatta gataccctgg tagtccacgc tgtaaacgct gtcgatttgg gggttggggt 840
ttaactctgg cacccgtagc taacgtgata aatcgaccgc ctggggagta cggccgcaag 900
gttaaaactc aaatgaattg acgggggccn gcacaagcgg tggagcatgt ggtttaattc 960
gatgcaacgc gaagaacctt acctactctt gacatcctaa gaagagctca gagatgagct 1020
tgtgccttcg ggaacttaga gacaggtgct gcatggctgt cgtcagctcg tgttgtgaaa 1080
tgttgggtta agtcccgcaa cgagcgcaac ccttatcctt tgttgccagc gacttggtcg 1140
ggaactcaaa ggagactgcc agtgataaac tggaggaagg tnggggatgac gtcaagtcat 1200
catggccctt acgagtaggg ctacacacgt gctacaatgg cgtatacaga gggaagcgaa 1260
gctgcgaggt ggagcgaatc tcataaagta cgtctaagtc cggattggag tctgcaactc 1320
gactccatga agtcggaatc gctagtaatc gcgaatcaga atgtcgcggt gaatacgttc 1380
ccgggcnttg tacacaccgc ccgtcacacc atgggagtgg gttgtaccag aagtagatag 1440
cttaaccttt tggagggcgt ttaccacggt atgattcatg actgggg 1487

<210> 38
<211> 1532
<212> DNA
<213> Bordetella bronchiseptica

<400> 38
tgaactgaag agtttgatcc tggctcagat tgaacgctgg cgggatgctt tacacatgca 60
agtcggacgg cagcacgggc ttcggcctgg tggcgagtgg cgaacgggtg agtaatgtat 120
cggaacgtgc ccagtagcgg gggataacta cgcgaaagcg tggctaatac cgcatacgcc 180

ctacgggga aagcgggga ccttcgggcc tcgcactatt ggagcggccg atatcggatt 240
agctagttgg tgggg taacg gcctaccaag gcgacgatcc gtagctggtt tgagaggacg 300
accagccaca ctggga ctga gacacggccc agactcctac gggaggcagc agtggggaat 360
tttggacaat gggggcaacc ctgatccagc catcccgcgt gtgcgatgaa ggccttcggg 420
ttgtaaagca cttttggcag gaaagaaacg gcacgggcta atatcctgtg caactgacgg 480
tacctgcaga ataagcaccg gctaactacg tgccagcagc cgcggtaata cgtagggtgc 540
aagcgttaat cggaattact gggcgtaaag cgtgcgcagg cggttcggaa agaaagatgt 600
gaaatcccag ggcttaacct tggaactgca tttttaacta ccgggctaga gtgtgtcaga 660
gggaggtgga attccgcgtg tagcagtgaa atgcgtagat atgcggagga acaccgatgg 720
cgaaggcagc ctcctgggat aacactgacg ctcatgcacg aaagcgtggg gagcaaacag 780
gattagatac cctggtagtc cacgccctaa acgatgtcaa ctagctgttg gggccttcgg 840
gccttggtag cgcagctaac gcgtgaagtt gaccgcctgg ggagtacggt cgcaagatta 900
aaactcaaag gaattgacgg ggacccgcac aagcggtgga tgatgtggat taattcgatg 960
caacgcgaaa aaccttacct acccttgaca tgtctggaat cccgaagaga tttgggagtg 1020
ctcgcaagag aaccggaaca caggtgctgc atggctgtcg tcagctcgtg tcgtgagatg 1080
ttgggttaag tcccgcaacg agcgcaaccc ttgtcattag ttgctacgaa agggcactct 1140
aatgagactg ccggtgacaa accggaggaa ggtggggatg acgtcaagtc ctcatggccc 1200
ttatgggtag ggcttcacac gtcatacaat ggtcgggaca gaagggtcgcc aacccgcgag 1260
ggggagccaa tcccagaaac ccgatcgtag tccggatcgc agtctgcaac tcgactgcgt 1320
gaagtcggaa tcgctagtaa tcgcggatca gcatgtcgcg gtgaatacgt tcccgggtct 1380
tgtacacacc gcccgtcaca ccatgggagt gggttttacc agaagtagtt agcctaaccg 1440
caaggggggc gattaccacg gtaggattca tgactggggt gaagtcgtaa caaggtagcc 1500
gtatcggaag gtgcggctgg atcacctcct tt 1532

<210> 39
<211> 1485
<212> DNA
<213> Bordetella parapertussis

<400> 39
attgaacgct ggcgggatgc tttacacatg caagtcggac ggcagcacgg gcttcggcct 60
ggtggcgagt ggcgaacggg tgagtaatgt atcggaacgt gcccagtagc gcgggataac 120
tacgcgaaag cgtggctaat accgcatacg ccctacgggg gaaagcgggg gactttcggg 180
cctcgcacta ttggagcggc cgatatcgga ttagctagtt ggtgaggtaa cggcctacca 240
aggcgacgat ccgtagctgg tttgagagga cgaccagcca cactgggact gagacacggc 300
ccagactcct acgggaggca gcagtgggga attttggaca atgggggcaa ccctgatcca 360
gccatcccgc gtgtgcgatg aaggccttcg ggttgtaaag cacttttggc aggaaagaaa 420
cggcacgggc taatatcctg tgcaactgac ggtacctgca gaataagcac cggctaacta 480
cgtgccagca gccgcggtaa tacgtagggt gcaagcgtta atcggaatta ctgggcgtaa 540
agcgtgcgca ggcggttcgg aaagaaagat gtgaaatccc agggcttaac cttggaactg 600
catttttaac taccgggcta gagtgtgtca gagggaggtg gaattccgcg tgtagcagtg 660
aaatgcgtag atatgcggag gaacaccgat ggcgaaggca gcctcctggg ataacactga 720
cgctcatgca cgaaagcgtg gggagcaaac aggattagat accctggtag tccacgccct 780
aaacgatgtc aactagctgt tggggccttc gggccttggt agcgcagcta acgcgtgaag 840
ttgaccgcct ggggagtacg gtcgcaagat taaaactcaa aggaattgac ggggacccgc 900
acaagcggtg gatgatgtgg attaattcga tgcaacgcga aaaaccttac ctacccttga 960
catgtctgga atcccgaaga gatttgggag tgctcgcaag agaaccggaa cacaggtgct 1020
gcatggctgt cgtcagctcg tgtcgtgaga tgttgggtta agtcccgcaa cgagcgcaac 1080
ccttgtcatt agttgctacg aaagggcact ctaatgagac tgccggttac aaaccggagg 1140
aaggtgggga tgacgtcaag tcctcatggc ccttatgggt agggcttcac acgtcatacа 1200
atggtcggga cagagggtcg ccaacccgcg aggggagcc aatcccagaa acccgatcgt 1260
agtccggatc gcagtctgca actcgactgc gtgaagtcgg aatcgctagt aatcgcggat 1320
cagcatgtcg cggtgaatac gttcccgggt cttgtacaca ccgcccgtca caccatggga 1380
gtgggtttta ccagaagtag ttagcctaac cgcaaggggg gggcgattac cacggtagga 1440
ttcatgactg gggtgaagtc gtaacaaggt agccgtatcg gaagg 1485

<210> 40
<211> 1464
<212> DNA
<213> Bordetella pertussis

<220>
<221> modified_base
<222> (87)..(1391)
<223> N = A, C, G or T/U

<400> 40
aactgaagag tttgatcctg gctcagattg aacgctggcg ggatgcttta cacatgcaag 60
tcggacggca gcacgggctt cggcctngtg gcgagtggcg aacggtgtag taatgtatcg 120
gaacgtgccc agtagcgggg gataactacg cgaaagcgta gctaataccg catacgccct 180
acgggggaaa gcgggggacc ttcgggcctc gcactattgg agcggccgat atcggattag 240
ctngttggtg gggtaacggc ctaccaaggc gacgatccgt agctggtttg agaggacgac 300
cagccacact gggactgaga cacggcccag nctcctacgg gaggcagcag tggggaattt 360
tggacaatgg gggcaaccct gatccagcca tccgcgtgt gcgatgaagg ccttcgggtt 420
gtaaagcact tttggcagga aagaaacggc acgggctaat atcctgtgca actgacggta 480
cctgcagaat aagcaccggc taactacgtg ccagcagccg cggtaatacg tagggtgcaa 540
gcgttaatcg gaattactgg gcgtaaagcg tgcgcaggcg gttcggaaag aaagatgtga 600
aatcccaggg cttaaccttg gaactgcatt tttaactacc gggctagagt gtgtcagagg 660
gaggtggaat tccgcgtgta gcagtgaaat gcgtagatat gcggaggaac accgatggcg 720
aaggcagcct cctgggataa cactgacgct catgcacgaa agtgtgggga gcaaacagga 780
ttagataccc tggtagtcca cgccctaaac gatgtcaact agctgttggg gccttcgggc 840
cttggtagcg cagctaacgc gtgaagttga ccgcctgggg agtacggtcg caagattaaa 900
actcaaagga attgacgggg acccgcacaa gcggtggatg atgtggatta attcgatgca 960
acgcgaaaaa ccttacctac ccttgacatg tctggaatcc cgaagagatt tgggagtgct 1020
cgcaagagaa ccggaacaca ggtgctgcat ggctgtcgtc agctcgtgtc gtgagatgtt 1080
gggttaagtc ccgcaacgag cgcaaccctt gtcattagtt gctacgaaag ggcactctaa 1140
tgagactgcc ggtgacaaac cggaggaagg tggggatgac gtgaagtcct catggccctt 1200
atgggtaggg cttcacacgt catacaatgg tcgggacaga gggttgncaa ccgcgaggg 1260
ggagccaatc ccagaaaccc ggtcgtngtc cggatcgcag tctgcaactc gactgcgtga 1320
agtcggaatc gctagtaatc gcggatcagc atgtcgcggt gaatacgttc ccgggtcttg 1380
tacacaccgc ncgtcacacc atgggagtgg gttttaccag aagtagttag cctaaccgca 1440
agggggcga ttaccacggt agga 1464

<210> 41
<211> 1535
<212> DNA
<213> Burkholderia cepacia

<400> 41
taaactgaag agtttgatcc tggctcagat tgaacgctgg cggcatgctt aacacatgca 60
agtcgaacgg cagcacgggt gcttgcacct ggtggcgagt ggcgaacggg tgagtaatac 120
atcggaacat gtcctgtagt gggggatagc ccggcgaaag ccggattaat accgcatacg 180
atctacggat gaaagcgggg gaccttcggg cctcgcgcta tagggttggc gatggctgat 240
tagctagttg gtggggtaaa ggcctaccaa ggcgacgatc agtagctggt ctgagaggac 300
gaccagccac actgggactg agacacggcc cagactccta cgggaggcag cagtggggaa 360
ttttggacaa tgggcgaaag cctgatccag caatgccgcg tgtgtgaaga aggccttcgg 420
gttgtaaagc acttttgtcc ggaaagaaat ccctggctct aatacagtcg ggggatgacg 480
gtaccggaag aataagcacc ggctaactac gtgccagcag ccgcggtaat acgtagggtg 540
caagcgttaa tcggaattac tgggcgtaaa gcgtgcgcag gcggtttgct aagaccgatg 600

tgaaatcccc gggctcaacc tgggaactgc attggtgact ggcaggctag agtatggcag 660
agggggggtag aattccacgt gtagcagtga aatgcgtaga gatgtggagg aataccgatg 720
gcgaaggcag ccccctgggc caatactgac gctcatgcac gaaagcgtgg ggagcaaaca 780
ggattagata ccctggtagt ccacgcccta aacgatgtca actagttgtt ggggattcat 840
ttccttagta acgtagctaa cgcgtgaagt tgaccgcctg gggagtacgg tcgcaagatt 900
aaaactcaaa ggaattgacg gggacccgca caagcggtgg atgatgtgga ttaattcgat 960
gcaacgcgaa aaaccttacc taccccttgac atggtcggaa tcctgctgag aggtgggagt 1020
gctcgaaaga gaaccggcgc acaggtgctg catggctgtc gtcagctcgt gtcgtgagat 1080
gttgggttaa gtcccgcaac gagcgcaacc cttgtcctta gttgctacgc aagagcactc 1140
taaggagact gccggtgaca aaccggagga aggtggggat gacgtcaagt cctcatggcc 1200
cttatgggta gggcttcaca cgtcatacaa tggtcggaac agagggttgc caacccgcga 1260
gggggagcta atcccagaaa acccatcgta gtccggattg cactctgcaa ctcgagtgca 1320
tgaagctgga atcgctagta atcgcggatc agcatgccgc ggtgaatacg ttcccgggtc 1380
ttgtacacac cgcccgtcac accatgggag tgggtttcac cagaagtggc tagtctaacc 1440
gcaaggagga cggtcaccac ggtaggattc atgactgggg tgaagtcgta acaaggtagc 1500
cgtatcggaa ggtgcggctg gatcacctcc tttct 1535

<210> 42
<211> 1488
<212> DNA
<213> Burkholderia mallei

<400> 42
agattgaacg ctggcggcat gccttacaca tgcaagtcga acggcagcac gggcttcggc 60
ctggtggcga gtggtgaacg ggtgagtaat acatcggaac atgtcctgta gtggggata 120
gcccggcgaa agccggatta ataccgcata cgatctgagg atgaaagcgg gggaccttcg 180
ggcctcgcgc tatagggttg gccgatggct gattagctag ttggtggggt aaaggcctac 240
caaggcgacg atcagtagct ggtctgagag gacgaccagc cacactggga ctgagacacg 300
gcccagactc ctacgggagg cagcagtggg gaattttgga caatggggcgc aagcctgatc 360
cagcaatgcc gcgtgtgtga agaaggcctt cgggttgtaa agcacttttg tccggaaaga 420
aatcattctg gctaataccc ggagtggatg acggtaccgg aagaataagc accggctaac 480
tacgtgccag cagccgcggt aatacgtagg gtgcgagcgt taattggaat tactggcgt 540
aaagcgtgcg caggcggttt gctaagaccg atgtgaaatc cccgggctca acctgggaac 600
tgcattggtg actggcaggc tagagtatgg cagaggggg tagaattcca cgtgtagcag 660
tgaaatgcgt agagatgtgg aggaataccg atggcgaagg cagccccctg ggccaatact 720
gacgctcatg cacgaaagcg tggggagcaa acaggattag ataccctggt agtccacgcc 780
ctaaacgatg tcaactagtt gttggggatt catttcctta gtaacgtagc taacgcgtga 840
agttgaccgc ctggggagta cggtcgcaag attaaaactc aaaggaattg acggggaccc 900
gcacaagcgg tggatgatgt ggattaattc gatgcaacgc gaaaaacctt acctaccctt 960
gacatggtcg gaagcccgat gagagttggg cgtgctcgaa agagaaccgg cgcacaggtg 1020
ctgcatggct gtcgtcagct cgtgtcgtga gatgttgggt taagtcccgc aacgagcgca 1080
acccttgtcc ttagttgcta cgcaagagca ctctaaggag actgccggtg acaaaccgga 1140
ggaaggtggg gatgacgtca agtcctcatg gcccttatgg gtagggcttc acacgtcata 1200
caatggtcgg aacagagggt cgccaacccg cgaggggcag ccaatcccag aaaaccgatc 1260
gtagtccgga ttgcactctg caactcgagt gcatgaagct ggaatcgcta gtaatcgcgg 1320
atcagcatgc cgcggtgaat acgttcccgg gtcttgtaca caccgcccgt cacaccatgg 1380
gagtgggttt taccagaagt ggctagtcta accgcaagga ggacggtcac cacggtagga 1440
ttcatgactg gggtgaagtc gtaacaaggt agccgtatcg gaaggtgc 1488

<210> 43
<211> 1610
<212> DNA
<213> Burkholderia pseudomallei

<400> 43
tctagatgcg tgctcgagcg gccgcccagt gctgcatgga tatctgctga attcggcttg 60
agcagtttga tcctggctca gattgaacgc tggcggcatg ccttacacat gcaagtcgaa 120
cggcagcacg ggcttcggcc tggtggcgag tggcgaacgg gtgagtaata catcggagca 180
tgtcctgtag tggggatag cccgggcgaaa gccgaatcaa taccgcatac gatctgagga 240
tgaaagcggg ggaccttcgg gcctcgcgct atagggtgtgg ccgatggctg attagctagt 300
tggtggggta aaggcctacc aaggcgacga tcagtagctg gtctgagagg acgaccagcc 360
acactgggac tgagacacgg cccagactcc tacgggaggc agcagtgggg aattttggac 420
aatgggcgca agcctgatcc agcaatgccg cgtgtgtgaa gaaggccttc gggttgtaaa 480
gcactttgt ccggaaagaa atcattctgg ctaataccgg gagtggatga cggtaccgaa 540
agaataagca ccggctaact acgtgccagc agccgcggta atacgtaggg tgcgagcgtt 600
aatcggggatt actgggcgta aagcgtgcgc aggcggtttg ctaagaccga tgtgaaatcc 660
ccgggctcaa cctgggaact gcattggtga ctggcagget agagtatggc agaggggggt 720
agaattccac gtgtagcagt gaaatgcgta gagatgtgga ggaataccga tggcgaaggc 780
agcccctgg gccaatactg acgctcatgc acgaaagcgt ggggagaaaa caggattaga 840
taccctggta gtccacgccc taaacgatgt caactagttg ttggggattc atttccttag 900
taacgtagct aacgcgcgaa gttgaccgcc tggggagtac ggtcgcaaga ttaaaactca 960
aaggaattga cggggacccg cacaagcggt ggatgatgtg gattaattcg atgcaacgcg 1020
aaaaacctta cctacccttg acatggtcgg aagcccgagt agagttgggc gtgctcgaaa 1080
gagaaccggc gcacaggtgc tgcatggctg tcgtcagctc gtgtcgtgag atgttgggtt 1140
aagtcccgca acgagcgcaa cccttgtcct tagttgctac gcaagagcac tctaaggaga 1200
ctgccggtga caaaccggag gaaggtgggg atgacgtcaa gtcctcatgg cccttatggg 1260
tagggcttca cacgtcatac aatggtcgga acagagggtc gccaacccgc gaggggagc 1320
caatcccaga aaaccgatcg tagtccggat tgcactctgc aactcgagtg catgaagctg 1380
gaatcgctag taatcgcgga tcagcatgcc gcggtgaata cgttcccggg tcttgtacac 1440
accgcccgtc acaccatggg agtgggtttt accagaagtg gctagtctaa ccgcaaggag 1500
gacggtcacc acggtaggat tcatgactgg ggtgaagtcg taacaaggta gccgtagaag 1560
ccgaattcca gcacactggc ggccgttact actggatccg agctcgtacc 1610

<210> 44
<211> 1544
<212> DNA
<213> Neisseria gonorrhoeae

<400> 44
tgaacataag agtttgatcc tggctcagat tgaacgctgg cggcatgctt tacacatgca 60
agtcggacgg cagcacaggg aagcttgctt ctcgggtggc gagtggcgaa cgggtgagta 120
acatatcgga acgtaccggg tagcggggga taactgatcg aaagatcagc taataccgca 180
tacgtcttga gagggaaagc aggggacctt cgggccttgc gctatccgag cggccgatat 240
ctgattagct ggttggcggg gtaaaggccc accaaggcga cgatcagtag cgggtctgag 300
aggatgatcc gccacactgg gactgagaca cggcccagac tcctacggga ggcagcagtg 360
gggaattttg gacaatgggc gcaagcctga tccagccatg ccgcgtgtct gaagaaggcc 420
ttcgggttgt aaaggacttt tgtcagggaa gaaaaggctg ttgccaatat cggcggccga 480
tgacggtacc tgaagaataa gcaccggcta actacgtgcc agcagccgcg gtaatacgta 540
gggtgcgagc gttaatcgga attactgggc gtaaagcggg cgcagacggt tacttaagca 600
ggatgtgaaa tccccgggct caacccggga actgcgttct gaactgggtg actcgagtgt 660
gtcagaggga ggtggaattc cacgtgtagc agtgaaatgc gtagagatgt ggaggaatac 720
cgatggcgaa ggcagcctcc tgggataaca ctgacgttca tgtccgaaag cgtgggtagc 780
aaacaggatt agataccctg gtagtccacg ccctaaacga tgtcaattag ctgttgggca 840
acttgattgc ttggtagcgt agctaacgcg tgaaattgac cgcctgggga gtacggtcgc 900
aagattaaaa ctcaaaggaa ttgacggggg cccgcacaag cggtggatga tgtggattaa 960
ttcgatgcaa cgcgaagaac cttacctggt tttgacatgt gcggaatcct ccggagacgg 1020
aggagtgcct tcgggagccg taacacaggt gctgcatggc tgtcgtcagc tcgtgtcgtg 1080
agatgttggg ttaagtcccg caacgagcgc aacccttgtc attagttgcc atcattcggt 1140
tgggcactct aatgagactg ccggtgacaa gccggaggaa ggtggggatg acgtcaagtc 1200

ctcatggccc ttatgaccag ggcttcacac gtcatacaat ggtcggtaca gagggtagcc 1260
aagccgcgag gcggagccaa tctcacaaaa ccgatcgtag tccggattgc actctgcaac 1320
tcgagtgcat gaagtcggaa tcgctagtaa tcgcagctca gcatactgcg gtgaatacgt 1380
tcccgggtct tgtacacacc gcccgtcaca ccatgggagt gggggatacc agaagtaggt 1440
agggtaaccg caaggagtcc gcttaccacg gtatgcttca tgactggggt gaagtcgtaa 1500
caaggtagcc gtaggggaac ctgcggctgg atcacctcct ttct 1544

<210> 45
<211> 1544
<212> DNA
<213> Neisseria meningitidis

<400> 45
tgaacataag agtttgatcc tggctcagat tgaacgctgg cggcatgctt tacacatgca 60
agtcggacgg cagcacagag aagcttgctt ctcgggtggc gagtggcgaa cgggtgagta 120
acatatcgga acgtaccgag tagtggggga taactgatcg aaagatcagc taataccgca 180
tacgtcttga gagagaaagc aggggacctt cgggccttgc gctattcgag cggccgatat 240
ctgattagct agttggtggg gtaaaggcct accaaggcga cgatcagtag cgggtctgag 300
aggatgatcc gccacactgg gactgagaca cggcccagac tcctacggga ggcagcagtg 360
gggaattttg gacaatgggc gcaagcctga tccagccatg ccgcgtgtct gaagaaggcc 420
ttcgggttgt aaaggacttt tgtcagggaa gaaaaggctg ttgctaatat cagcggctga 480
tgacggtacc tgaagaataa gcaccggcta actacgtgcc agcagccgcg gtaatacgta 540
gggtgcgagc gttaatcgga attactgggc gtaaagcggg cgcagacggt tacttaagca 600
ggatgtgaaa tccccgggct caacccggga actgcgttct gaactgggtg actcgagtgt 660
gtcagaggga ggtagaattc cacgtgtagc agtgaaatgc gtagagatgt ggaggaatac 720
cgatggcgaa ggcagcctcc tgggacaaca ctgacgttca tgcccgaaag cgtgggtagc 780
aaacaggatt agataccctg gtagtccacg ccctaaacga tgtcaattag ctgttgggca 840
acctgattgc ttggtagcgt agctaacgcg tgaaattgac cgcctgggga gtacggtcgc 900
aagattaaaa ctcaaaggaa ttgacgggga cccgcacaag cggtggatga tgtggattaa 960
ttcgatgcaa cgcgaagaac cttacctggt cttgacatgt acggaatcct ccggagacgg 1020
aggagtgcct tcgggagccg taacacaggt gctgcatggc tgtcgtcagc tcgtgtcgtg 1080
agatgttggg ttaagtcccg caacgagcgc aaccctttgtc attagttgcc atcattcagt 1140
tgggcactct aatgagactg ccggtgacaa gccggaggaa ggtggggatg acgtcaagtc 1200
ctcatggccc ttatgaccag ggcttcacac gtcatacaat ggtcggtaca gagggtagcc 1260
aagccgcgag gcggagccaa tctcacaaaa ccgatcgtag tccggattgc actctgcaac 1320
tcgagtgcat gaagtcggaa tcgctagtaa tcgcaggtca gcatactgcg gtgaatacgt 1380
tcccgggtct tgtacacacc gcccgtcaca ccatgggagt gggggatacc agaagtaggt 1440
aggataacca caaggagtcc gcttaccacg gtatgcttca tgactggggt gaagtcgtaa 1500
caaggtagcc gtaggggaac ctgcggctgg atcacctcct ttct 1544

<210> 46
<211> 1537
<212> DNA
<213> Pseudomonas aeruginosa

<400> 46
gaactgaaga gtttgatcat ggctcagatt gaacgctggc agcaggggcc ttcaacacat 60
gcaagtcgag cttatgaagg gagcttgcct tggattcagc ggcggacggg tgagtaatgc 120
ctaggaatct gcctggtagt gggggataac gtccggaaac ggccgctaat accgcataacg 180
tcctgaggga gaaagtcggg gatcttcgga cctcacgcta tcagatgagc ctaggtcgga 240
ttagctagtt ggtggggtaa aggcctacca aggcgacgat ccgtaactgg tctgagagga 300
tgatcagtca cactggaact gagacacggt ccagactcct acgggaggca gcagtgggga 360
atattggaca atgggcgcaa gcctgatcca gccatgccgc gtgtgtgaag aaggtcttcg 420
gattgtaaag cactttaagt tgggaggaag ggcagtaagt taataccttg ctgtttgacg 480

ttaccaacag aataagcacc ggctaacttc gtgccagcag ccgcggtaat acgaagggtg 540
caagcgttaa tcggaattac tgggcgtaaa gcgcgcgtaa gtggttcagc aagcttgatg 600
tgaaatcccc gggctcaacc tgggaactgc atccaaaagc tactgagcta gagtacggta 660
gaggtggtag aatttcctgt gtagcggtga aatgcgtaga tataggaagg aacaccagtg 720
gcgaaggcga ccacctggac tgtactgaca ctgaggtgcg aaagcgtggg gagcaaacag 780
gattagatac cctggtagtc cacgccgtaa acgatgtcga ctagccgttg gggtccttga 840
gatcttagtg gcgcacgtaa cgcgataagt cgaccgcctg gggagtacgg ccgcaaggtt 900
aaaactcaaa tgaattgacg gggcccgca caagcggtgg agcatgtggt ttaattcgaa 960
gcaacgcgaa gaaccttacc tggccttgac atgctgagaa ctttccagag atggattggt 1020
gccttcggga acagagacac aggtgctgca tggctgtcgt cagctcgtgt cgtgagatgt 1080
tgggttaagt cccgtaacga gcgcaaccct tgtccttagt taccagcacc tcgggtgggc 1140
actctaagga gactgccggt gacaaaccgg aggaaggtgg ggatgacgtc aagtcatcat 1200
ggcccttacg gccagggcta cacacgtgct acaatggtcg gtacaaaggg ttgccaagcc 1260
gcgagtggga gctaatccca taaaaccgat cgtagtccgg atcgcagtct gcaactcgac 1320
tgcgtgaagt cggaatcgct agtaatcgtg aatcagaatg tcacggtgaa tacgtcccct 1380
ggccttgtac acaccgcccg tcacaccatg ggagtgggtt gctccagaag tagctagtct 1440
aaccgcaagg ggacggtta ccacggagtg attcatgact ggggtgaagt cgtaacaagg 1500
tagccgtagg ggaacctgcg gctggatcac ctcctta 1537

<210> 47
<211> 1467
<212> DNA
<213> Vibrio cholerae

<220>
<221> modified_base
<222> (928)..(1464)
<223> N = A, C, G or T/U

<400> 47
attgaagagt ttgatcctgg ctcagattga acgctggcgg caggcctaac acatgcaagt 60
cgagcggcag cacagaggaa cttgttcctt gggtggcgag cggcggacgg gtgagtaatg 120
cctgggaaat tgcccggtag aggggataa ccattggaaa cgatggctaa taccgcataa 180
cctcgcaaga gcaaagcagg ggaccttcgg gccttgcgct accggatatg cccaggtggg 240
attagctagt tggtgaggta agggctcacc aaggcgacga tccctagctg gtctgagagg 300
atgatcagcc acactggaac tgagacacgg tccagactcc tacgggaggc agcagtgggg 360
aatattgcac aatgggcgca agcctgatgc agccatgccg cgtgtatgaa gaaggccttc 420
gggttgtaaa gtactttcag tagggaggaa ggtggttaag ttaatacctt aatcatttga 480
cgttacctac agaagaagca ccggctaact ccgtgccagc agccgcggta atacggaggg 540
tgcaagcgtt aatcggaatt actgggcgta aagcgcatgc aggtggtttg ttaagtcaga 600
tgtgaaagcc ctgggctcaa cctaggaatc gcatttgaaa ctgacaagct agagtactgt 660
agaggggggt agaatttcag gtgtagcggt gaaatgcgta gagatctgaa ggaataccgg 720
tggcgaaggc ggccccctgg acagatactg acactcagat gcgaaagcgt ggggagcaaa 780
caggattaga taccctggta gtccacgccg taaacgatgt ctacttggag gttgtgccct 840
agagtcgtgg ctttcggagc taacgcgtta agtagaccgc ctggggagta cggtcgcaag 900
attaaaactc aaatgaattg acgggggncc gcacaagcgg tggagcatgt ggtttaattc 960
gannca acgc gaagaacctt acctactctt gacatccaga gaatctagcg gagacgctgg 1020
agtgccttcg ggagctctga gacaggtgct gcatggctgt cgtcagctcg tgttgtgaaa 1080
tgttgggtta agtcccgcaa cgagcgcaac ccttatcctt gtttgccagc acgtaatggt 1140
gggaactcca gggagactgc cggtgataaa ccggaggaag gtggggacga cgtcaagtca 1200
tcatggccct tacgagtagg gctacacacg tgctacaatg gcgtatacag agggcagcga 1260
taccgcgagg tggagcgaat ctcacaaagt acgtcgtagt ccggattgga gtctgcaact 1320
cgactccatg aagtcggaat cgctagtaat cgcaaatcag aatgttgcgg tgaatacgtt 1380
cccgggcctt gtacacaccg cccgtcacac catgggagtg ggctgcaaaa gaagcangta 1440
gtttaacctt cgggaggacg cttnccc 1467

<210> 48
<211> 1485
<212> DNA
<213> Yersinia enterocolitica

<220>
<221> modified_base
<222> (1)..(1484)
<223> N = A, C, G or T/U

<400> 48
naattgaaga gtttgatcat ggctcagatn gaacgctggc ggcaggccta acacatgcaa 60
gtcgagcggc agcggggaagn agtttactac tttcngggcg agcggcgnac gggtgagtaa 120
tgtctgggaa actgcctgat ggagggggat aactactgga aacggtagct aataccgcat 180
aacgtcttcg gaccaaagtg ggggacctta gggcctcacg ccatcngatg tgcccagatg 240
ggattagcta gtaggtgggg taatggctca cctaggcgac gatccctagc tggtctgaga 300
ggatgaccag ccacactgga actgagacac ggtccagact cctacgggag gcagcagtgg 360
ggaatattgc acaatgggcg caagcctgat gcagccatgc cgcgtgtgtg aagaaggcct 420
tcgggttgta aagcactttc agcgaggagg aaggccaata acttaatacg ttgttggatt 480
gacgttactc gcagaagaag caccggctaa ctccgtgcca gcagccgcgg taatacggag 540
ggtgcaagcg ttaatcggaa ttactgggcg taaagcgcac gcaggcggtt tgttaagtca 600
gatgtgaaat ccccgcgctt aacgtgggna cngcatttga aactggcaag ctagagtctt 660
gtagaggggg gtagaattcc aggtgtagcg gtgaaatgcg tagagatctg naggaatacc 720
ggtggcgaag gcggcccccc ggacaaagac tgacgctcag gtgcgaaagc gtggggagca 780
aacaggatta gataccctgg tagtccacgc tgtaaacgat gtcgacttgg aggttgtgcc 840
cttgaggcgt ggcttccgga gctaacgcgt taagtcgacc gcctggggag tacggccgca 900
aggttaaaac tcaaatgaat tnncgggggc cngcacaagc ggtggagcat gtggtttaat 960
tcgatgcaac gcgaagaacc ttacctactc ttgacatcca cggaatttag cagagatgct 1020
ttagtgnctt cgggaaccgt gagacaggtg ctgcatggct gtcgtcagct cgtgttgtga 1080
aatgttgggt taagtcccgc aacgagcgca acccttatcc tttgttgcca gcacgtaatg 1140
gtgggaactc aaaggagact gccggtgata aaccggagga aggtggggat gacgtcaagt 1200
catcatggcc cttacgagta gggctacaca cgtgctacaa tggcagatac aaagtgaagc 1260
gaactcgcga gagcaagcgg accacataaa gtctgtcgta gtccggattg gagtctgcaa 1320
ctcgactcca tgaagtcgga atcgctagta atcgtagatc agaatgctac ggtgaatacg 1380
ttcccgggcc ttgtacacac cgcccgtcac accntgggag tgggttgcaa aagaagtagg 1440
tagcttaacn ttcgggaggg cgcgtaccac tttgtgattc nngnc 1485

<210> 49
<211> 2927
<212> DNA
<213> Bacillus subtilis

<400> 49
ggttaagtta gaaagggcgc acggtggatg ccttggcact aggagccgat gaaggacggg 60
acgaacaccg atatgcttcg gggagctgta agcaagcttt gatccggaga tttccgaatg 120
gggaaaccca ccactcgtaa tggagtggta tccatatctg aattcatagg atatgagaag 180
gcagacccgg ggaactgaaa catctaagta cccggagaag agaaagcaaa tgcgattccc 240
tgagtagcgg cgacgaacac gggatcagcc caaaccaaga ggcttgcctc tgtggttgta 300
ggacactctg tacggagtta caaaagaacg aggtagatga agaggtctgg aaagggcccg 360
ccataggagg taacagccct gtagtcaaaa cttcgttctc tcctgagtgg atcctgagta 420
cggcggaaca cgtgaaattc cgtcggaatc cgggaggacc atctcccaag gctaaatact 480
ccctagtgac cgatagtgaa ccagtaccgt gagggaaagg tgaaaagcac cccggaaggg 540
gagtgaaaga gatcctgaaa ccgtgtgcct acaagtagtc agagcccgtt aacggtgatg 600

gcgtgccttt tgtagaatga accggcgagt tacgatccg tgcaaggtta agcagaagat 660
gcggagccgc agcgaaagcg agtctgaata gggcgcatga gtacgtggtc gtagacccga 720
aaccaggtga tctacccatg tccagggtga agttcaggta acactgaatg gaggcccgaa 780
cccacgcacg ttgaaaagtg cggggatgag gtgtgggtag gggtgaaatg ccaatcgaac 840
ctggagatag ctggttctct ccgaaatagc tttagggcta gcctcaaggt aagagtcttg 900
gaggtagagc actgattgga ctaggggccc tcaccgggtt accgaattca gtcaaactcc 960
gaatgccaat gacttatcct tgggagtcag actgcgagtg ataagatccg tagtcgaaag 1020
ggaaacagcc cagaccgcca gctaaggtcc caaagtatac gttaagtgga aaaggatgtg 1080
gagttgctta gacaaccagg atgttggctt agaagcagcc accatttaaa gagtgcgtaa 1140
tagctcactg gtcgagtgac tctggcgcga aaatgtaccg gggctaaacg tatcacggaa 1200
gctgcggact gttcttcgaa cagtggtagg agagcgttct aaaggctgtg aagccagacc 1260
ggaaggactg gtggacggct tagaagtgag aatgccggta tgagtagcga aaagaggggt 1320
gagatccact ccaccgaatg cctaagggtt cctgaggaag gctcgtccgc tcagggttag 1380
tcgggaccta agccgaggcc gaaaggcgta ggcgatggac aacaggttga tattcctgta 1440
ccacctcctc accatttgag caatgggggg tcgcaggagg atagggtaag cgcggtattg 1500
gatatccgcg tccaagcagt taggctggga aataggcaaa tccgtttccc ataaggctga 1560
gctgtgatgg cgagcgaaat atagtagcga agttcctgat tccacactgc caagaaaagc 1620
ctctagcgag gtgagaggtg cccgtaccgc aaaccgtcac aggtaggcga ggagagaatc 1680
ctaaggtgat cgagagaact ctcgttaagg aactcggcaa aatgaccccg taacttcggg 1740
agaaggggtg ctctgttagg gtgcaagccc gagagagccg cagtgaatag gcccaggcga 1800
ctgtttagca aaaacacagg tctctgcgaa gccgtaaggc gaagtatagg ggctgacgcc 1860
tgcccggtgc tggaaggtta agaggagcgc ttagcgtaag cgaaggtgcg aattgaagcc 1920
ccagtaaacg gcggccgtaa ctataacggt cctaaggtag cgaaattcct tgtcgggtaa 1980
gttccgaccc gcacgaaagg cgcaacgatc tgggcgctgt ctcaacgaga gactcggtga 2040
aattatagta cctgtgaaga tgcaggttac cgcgcacagg acggaaagac cccgtggagc 2100
tttactgcag cctgatattg aatgttggta cagcttgtac aggataggta ggagccttgg 2160
aaaccggagc gccagcttcg gtggaggcat cggtgggata ctacccctgg tgtattgacc 2220
ttctaacccc ccgcccttat cgggcgggga gacagtgtca ggtggggcagt ttgactgggg 2280
cggtcgcctc ctaaaaggta acggaggcgc ccaaaggttc cctcagaatg gttggaaatc 2340
attcgcagag tgtaaaggca caagggagct tgactgcgag acctacaagt cgagcaggga 2400
cgaaagtcgg gcttagtgat ccggtggttc cgcatggaag ggccatcgct caacggataa 2460
aagctacccc ggggataaca ggcttatctc ccccaagagc tccacatcga cgggggaggtt 2520
tggcacctcg atgtcggctc atcgcatcct ggggctgtag tcggtcccaa gggttgggct 2580
gttcgcccat taaagcggta cgcgagctgg gttcagaacg tcgtgagaca gttcggtccc 2640
tatccgtcgc gggcgctgga aatttgagag gagctgtcct tagtacgaga ggaccgggat 2700
ggacgcaccg ctggtgtacc agttgttctg ccaagggcat cgctgggtag ctatgtgcgg 2760
acgggataag tgctgaaagc atctaagcat gaagcccccc tcaagatgag atttccaatt 2820
ccgcaaggaa gtaagatccc tgaaagatga tcaggttgat aggtctgagg tggaagtgtg 2880
gcaacacatg gagctgacag atactaatcg atcgaggact taaccat 2927

<210> 50
<211> 2922
<212> DNA
<213> Bacillus anthracis

<400> 50
ggttaagtta gaaagggcgc acggtggatg ccttgacact aggagtcgat gaaggacggg 60
actaacgccg atatgcttcg gggagctgta agtaagcttt gatccgaaga tttccgaatg 120
gggaaaccca ccatacgtaa tggtatggta tccttatctg aatacatagg gtaaggaaga 180
cagacccagg gaactgaaac atctaagtac ctggaggaag agaaagcaaa tgcgatttcc 240
tgagtagcgg cgagcgaaac ggaacatagc ccaaaccaag aggcttgcct cttgggggttg 300
taggacattc tatacggagt tacaaaggaa cgaggtagac gaagcgacct ggaaaggtcc 360
gtcgtagagg gtaacaaccc cgtagtcgaa acttcgttct ctcttgaatg tatcctgagt 420
acggcggaac acgtgaaatt ccgtcggaat ctgggaggac catctcccaa ggctaaatac 480
tccctagtga tcgatagtga accagtaccg tgaggtaaag gtgaaaagca ccccggaagg 540

ggagtgaaag agatcctgaa accgtgtgcc tacaaatagt cagagcccgt taacgggtga 600
tggcgtgcct tttgtagaat gaaccggcga gttacgatcc cgtgcgaggt taagctgaag 660
aggcggagcc gcagcgaaag cgatctgaa tagggcgttt agtacgtggt cgtagacccg 720
aaaccaggtg atctacccat gtccaggtg aagttcaggt aacactgaat ggaggcccga 780
acccacgcac gttgaaaagt gcggggatga ggtgtgggta gcggagaaat tccaatcgaa 840
cctggagata gctggttctc cccgaaatag ctttaggggct agccttaagt gtaaggatct 900
tggaggtaga gcactgattg gactagcggt cctcatcgga ttaccgaatt cagtcaaact 960
ccgaatgcca atgacttatc cttaggagtc agactgcgag tgataagatc cgtagtcaaa 1020
agggaaacag cccagaccgc cagctaaggt cccaaagtgt gtattaagtg gaaaaggatg 1080
tggagttgct tagacaacta ggatgttggc ttagaagcag ccaccattta aagagtgcgt 1140
aatagctcac tagtcgagtg actctgcgcc gaaaatgtac cggggctaaa tacaccaccg 1200
aagctgcgga ttgataccaa tggtatcagt ggtaggggag cgttctaagg acagtgaagt 1260
cagaccggaa ggactggtgg agtgcttaga agtgagaatg ccggtatgag tagcgaaaga 1320
cgggtgagaa tcccgtccac cgaatgccta aggttcctg aggaaggctc gtccgctcag 1380
ggttagtcag gacctaagcc gaggccgaca ggcgtaggcg atggacaaca ggttgatatt 1440
cctgtaccac ctctttatcg tttgagcaat ggagggacgc agaaggatag aagaagcgtg 1500
cgattggttg tgcacgtcca agcagttagg ctgataagta ggcaaatccg cttatcgtga 1560
aggctgagct gtgatgggga agctccttat ggagcgaagt ctttgattcc ccgctgccaa 1620
gaaaagcttc tagcgagata aaaggtgcct gtaccgcaaa ccgacacagg taggcgagga 1680
gagaatccta aggtgtgcga gagaactctg gttaaggaac tcggcaaaat gaccccgtaa 1740
cttcgggaga aggggtgctt tcttaacgga aagccgcagt gaataggccc aagcgactgt 1800
ttagcaaaaa cacagctctc tgcgaagccg taaggcgaag tatagggggt gacacctgcc 1860
cggtgctgga aggttaagga gaggggttag cgtaagcgaa gctctgaact gaagccccag 1920
taaacggcgg ccgtaactat aacggtccta aggtagcgaa attccttgtc gggtaagttc 1980
cgaccgcgac gaaaggtgta acgatttggg cactgtctca accagagact cggtgaaatt 2040
atagtacctg tgaagatgca ggttacccgc gacaggacgg aaagaccccg tggagcttta 2100
ctgtagcctg atattgaatt ttggtacagt ttgtacagga taggcgggag cctttgaaac 2160
cggagcgcta gcttcggtgg aggcgctggt gggataccgc cctgactgta ttgaaattct 2220
aacctacggg tcttatcgac ccgggagaca gtgtcaggtg ggcagtttga ctggggcggt 2280
cgcctcctaa agtgtaacgg aggcgcccaa aggttccctc agaatggttg gaaatcattc 2340
gtagagtgca aaggcataag ggagcttgac tgcgagacct acaagtcgag cagggacgaa 2400
agtcgggctt agtgatccgg tggttccgca tggaagggcc atcgctcaac ggataaaagc 2460
taccccgggg ataacaggct tatctccccc aagagtccac atcgacgggg aggtttggca 2520
cctcgatgtc ggctcatcgc atcctggggc tgtagtcggt cccaaggggtt gggctgttcg 2580
cccattaaag cggtacgcga gctgggttca gaacgtcgtg agacagttcg gtccctatcc 2640
gtcgtgggcg taggaaattt gagaggagct gtccttagta cgagaggacc gggatggacg 2700
caccgctggt gtaccagttg ttctgccaag ggcatagctg ggtagctatg tgcggaaggg 2760
ataagtgctg aaagcatcta agcatgaagc ccccctcaag atgagatttc ccatagcgta 2820
agctagtaag atccctgaaa gatgatcagg ttgatagggt cgaggtggaa gcatggtgac 2880
atgtggagct gacgaatact aatagatcga ggacttaacc at 2922

<210> 51
<211> 2912
<212> DNA
<213> Enterococcus faecalis

<400> 51
ggttaagtga ataagggcgc acggtggatg ccttggcact aggagccgat gaaggacggg 60
actaacaccg atatgctttg gggagctgta agtaagctat gatccagaga tttccgaatg 120
ggggaaccca atatctttta taggatatta cttttcagtg aatacatagc tgattagagg 180
tagacgcaga gaactgaaac atcttagtac ctgcaggaag agaaagaaaa ttcgattccc 240
tgagtagcgg cgagcgaaac gggaagagcc caaaccaaca agcttgcttg ttgggtgttg 300
aggactccaa tatggtagtc tgttagtata gttgaaggat ttggaaaatt ccgctaaaga 360
gggtgaaagc cccgtagacg aaatgctaac aacacctagg aggatcctga gtacggcgga 420
acacgagaaa ttccgtcgga atccgcgggg accatcccgc aaggctaaat actcccctagt 480

gaccgatagt gaaccagtac cgtgagggaa aggtgaaaag caccccggaa ggggagtgaa 540
atagatcctg aaaccgtgtg cctacaacaa gtcaaagctc gttaatgagt gatggcgtgc 600
cttttgtaga atgaaccggc gagttacgat tgcatgcgag gttaagtcga agagacggag 660
ccgcagcgaa agcgagtctg aatagggcga atgagtatgt agtcgtagac ccgaaaccat 720
gtgatctacc catgtccagg ttgaaggtgc ggtaaaacgc actggaggac cgaacccacg 780
tacgttgaaa agtgcggggg tgaggtgtgg gtagcggaga aattccaaac gaacttggag 840
atagctggtt ctctccgaaa tagctttagg gctagcctcg gaattgagaa tgatggaggt 900
agagcactgt ttggactagg ggcccatctc gggttaccga attcagataa actccgaatg 960
ccattcattt atatccggga gtcagactgc gagtgataag atccgtagtc gaaagggaaa 1020
caggcccagac caccagctaa ggtcccaaaa tatatgttaa gtggaaaagg atgtgggggtt 1080
gcacagacaa ctaggatgtt ggcttagaag cagccaccat ttaaagagtg cgtaatagct 1140
cactagtcga gtgaccctgc gccgaaaatg taccggggct aaacatatta ccgaagctgt 1200
ggactacacc attaggtgta gtggtaggag agcgttctaa gggcgttgaa ggtcgatcgt 1260
gaggacggct ggagcgctta gaagtgagaa tgccggtatg agtagcgaaa gacaggtgag 1320
aatcctgtcc accgtatgac taaggtttcc tggggaaggc tcgtccgccc agggttagtc 1380
gggacctaag ccgaggccga taggcgtagg cgatggacaa caggttgata ttcctgtacc 1440
agttgttttt gtttgagcaa tggagggacg cagtaggcta aggaatgcat gcgattggaa 1500
gtgcatgtcc aagcaatgag tcttgagtag agttaaatgc tttactcttt aaggacaagt 1560
tgtgacgggg agcgaaataa tagtagcgaa gttcctgatg tcacactgcc aagaaaagct 1620
tctagtgaga aaacaactgc ccgtaccgta aaccgacaca ggtagtcgag gagagtatcc 1680
taaggtgagc gagcgaactc tcgttaagga actcggcaaa atgaccccgt aacttcggga 1740
gaaggggtgc tgacttcggt cagccgcagt gaataggccc aagcgactgt ttatcaaaaa 1800
cacaggtctc tgcaaaatcg taagatgaag tataggggct gacgcctgcc cggtgctgga 1860
aggttaagag gatgggttag cttcggcgaa gctcagaatt gaagccccag taaacggcgg 1920
ccgtaactat aacggtccta aggtagcgaa attccttgtc gggtaagttc cgacccgcac 1980
gaaaggcgta acgatttggg cactgttcta acgagagact cggtgaaatt ttagtacctg 2040
tgaagatgca ggttacccgc gacaggacgg aaagacccca tggagcttta ctgtagtttg 2100
atattgagtg tttgtaccac atgtacagga taggtaggag ccgatgagac cggaacgcta 2160
gtttcggagg aggcgctggt gggatactac ccttgtgtta tgaaccctct aacccgcacc 2220
actaatcgtg gtgggagaca gtgtcagatg ggcagtttga ctggggcggt cgcctcctaa 2280
aaggtaacgg aggcgcccaa aggttccctc agaatggttg gaaatcattc gaagagtgta 2340
aaggcagaag ggagcttgac tgcgagacct acaagtcgag cagggacgaa agtcgggctt 2400
agtgatccgg tggttccgca tggaagggcc atcgctcaac ggtaaaagct accctgggga 2460
taacaggctt atctccccca agagtccaca tcgacgggga ggtttggcac ctcgatgtcg 2520
gctcgtcgca tcctggggct gtagtcggtc ccaagggttg ggctgttcgc ccattaaagc 2580
ggcacgcgag ctgggttcag aacgtcgtga gacagttcgg tccctatccg tcgcgggcgt 2640
tggaaatttg agaggagctg tccttagtac gagaggaccg ggatggactt accgctggtg 2700
taccagttgt tctgccaagg gcattgctgg gtagctatgt agggaaggga taaacgctga 2760
aagcatctaa gtgtgaagcc cacctcaaga tgagatttcc catttcttta agaaagtaag 2820
acccctgaga gatgatcagg tagataggtt ggaagtggaa ggctagtgat agttggagcg 2880
gaccaatact aatcggtcga ggacttaacc aa 2912

<210> 52
<211> 2898
<212> DNA
<213> Lactococcus lactis

<400> 52
ggcaaagtta ataagggcgc acggtggatg ccttggcact aagagccgat gaaggacgtg 60
actaacgacg atattctagg gggagcagta agtacgcatt gatccctagg tctccgaatg 120
ggaaaaccca gctgctacta gcagttattc atgagtgaat acatagctca tgtaaaggta 180
acgcagagaa ctgaaacatc taagtacctg caggaagaga aagtaaaaac gatttcgtaa 240
gtagcggcga gcgaacgcga agaagggcaa accaagaagc ttgcttcttg gggttgtagg 300
actgcaacgt ggacttaagc attatagtcg aataacctgg gaaggttaat caaagagggt 360
aataatcccg tagacgaaat agcgcttata cctagcagta tcctgagtag ggctggacac 420

gcgaaatcca gtttgaatcc gggaggacca tctcccaacc ctaaatactc cttagtgacc 480
gatagtgaac cagtaccgtg agggaaaggt gaaaagaacc cgagagggga gtgaaatagc 540
acctgaaacc gtgtgcctac aagaagttcg agcccgttaa tgggtgagag cgtgcctttt 600
gtagaatgaa ccggcgagtt acgttatgat gcgaggttaa gttgaagaga cggagccgta 660
gggaaaccga gtctgaatag ggcgacttag tatcatgatg tagacccgaa acctagtgac 720
ctatccatga gcagggtgaa ggtgtggtaa gacgcactgg aggcccgaac caggacacgt 780
tgaaaagtgt ttggatgact tgtggatagc ggagaaattc caaacgaact gggagatagc 840
tggttctctc cgaaatagct ttagggctag cgtcgaaatg taagtgtatt ggaggtagag 900
cactgtttgg gtgaggggtc cgtctaggat taccaatctc agataaactc cgaatgctaa 960
tacacatgtt cggcagtcag actgcgagtg ctaagatccg tagtcgaaag ggaaacagcc 1020
cagaccaaca gctaaggtcc caaaatatat gttaagtgga aaaggatgtg gggttgcaca 1080
gacaactagg atgttagctc agaagcagct atcattcaaa gagtggctaa tagctcacta 1140
gtcgagtgac cctgcgccga aaatgtaccg gggctaaaca tattaccgaa gctttggatt 1200
gatattttat caatggtagg agagcgttct taaccgcgat gaaggtatac cgtgaggagt 1260
gctggagcgt taagaagtga gaatgccggt atgagtagcg caagataagt gagaatctca 1320
tccaccgtaa gactaaggtt tccaggggaa ggctcgtccg ccctgggtta gtcgggacct 1380
aaggcgaggc cgaaaggcgt agtcgatgga caactggttg atattccagt actagatatg 1440
atcgtgatgg agggacgcag taggctaaga gatgccagtt aatggattct ggtctaagca 1500
gtgaggtgtg agatgtgtca aatgcatttc tctttaacat tgagctgtga tggggaagca 1560
actacggttg cgaactctct gatgtcacac tgccaagaaa agcttctagc gtaaagtcat 1620
atctacccgt accgcaaacc gacacaggtg gtcgaggcga gtagcctcag gtgatcgaga 1680
gaactctcgt taaggaactc ggcaaaatag ccccgtaact tcgggagaag gggtgctggt 1740
gtaaaagcca gccgcagtga ataggcccaa gcaactgttt atcaaaaaca cagctctctg 1800
ctaaaccgca aggtgatgta tagggggtga cgcctgcccg gtgctggaag gttaagagga 1860
gtgcttagac gtaagtcgaa ggtatgaatt gaagccccag taaacggcgg ccgtaactat 1920
aacggtccta aggtagcgaa attccttgtc gggtaagttc cgaccgcac gaaaggcgta 1980
atgatttggg cactgtctca acgagagact cggtgaaatt ttagtacctg tgaagatgca 2040
ggttacccgc gacaggacgg aaagacccca tggagcttta ctgtagtttg atattgagta 2100
cctgtaagtc atgtacagga taggtaggag ccattgaaat agggacgcta gttttcattg 2160
aggcgttgtt gggatactac ccttgactta tggttactct aacccgctgg cataatcggc 2220
cagggagaca gtgtctgacg gacagtttga ctggggcggt cgctcctaaa gagtaacgga 2280
ggcgctcaaa ggttggctca gattggttgg aaatcaatcg tagagtgtaa aggtaaaagc 2340
cagcttgact gcgagagcta caactcgagc aggtaggaaa ctaggactta gtgatccggt 2400
ggtaccgcat ggaagggcca tcgctcaacg gataaaagct accctgggga taacaggctt 2460
atctccccca agagttcaca tcgacgggga ggtttggcac ctcgatgtcg gctcgtcgca 2520
tcctggggct gtagtcggtc ccaaggttg ggctgttcgc cattaaagcg gcacgcgagc 2580
tgggttcaga acgtcgtgag acagttcggt ccctatccgt cgcgggcgta ggtaatttga 2640
gaggatctgt ccttagtacg agaggaccgg gatggactta ccgctggtgt accagttgtt 2700
ccgccaggag cacggctgga tagctatgta gggaagggat aagcgctgaa agcatctaag 2760
tgcgaagccc acctcaagat gagattaccc attcgtaaga attaagagcc cagagagatg 2820
atctggtaga taggctggaa gtggaagagt tgcgagactt ggagcggacc agtactaatc 2880
gctcgaggac tttaccaa 2898

<210> 53
<211> 2932
<212> DNA
<213> Listeria monocytogenes

<400> 53
ggttaagtta gaaagggcgc acggtggatg ccttggcact aggagccgaa gaaggacggg 60
actaacaccg atatgctttg gggagctgta cgtaagcgtt gatccagaga tttccgaatg 120
ggggaaccca ctatctttag tcggatagta tccttacgtg aatacatagc gtgaggaagg 180
cagacccagg gaactgaaac atctaagtac ctggaggaag agaaagaaaa atcgatttcc 240
tgagtagcgg cgagcgaaac ggaaagagcc caaaccaaga agcttgcttc ttgggggttgt 300
aggacactct atacggagtt acaaaagaaa gttataaatg aagcggtctg gaaaggcccg 360

ccaaagacgg taacagcccg gtagttgaaa tggctttccc tccagagtgg atcctgagta 420
cggcggaaca cgtgaaattc cgtcggaatc cgggaggacc atctccccaag gctaaatact 480
ccctagtgac cgatagtgaa ccagtaccgt gagggaaagg tgaaaagcac cccggaaggg 540
gagtgaaaca gttcctgaaa ccgtgtgcct acaagtagtt agagcccgtt aatgggtgat 600
agcgtgcctt ttgtagaatg aaccggcgag ttacgatttg ttgcaaggtt aagcggaaaa 660
agcggagccg tagcgaaagc gagtctgaat agggcgcata agtaacaggt cgtagacccg 720
aaaccaggtg atctacccat gtccaggatg aaggtaaggt aatacttact ggaggtccga 780
acccacgcac gttgaaaagt gcggggatga ggtgtggtga gcggagaaat tccaatcgaa 840
cttggagata gctggttctc tccgaaatag ctttagggct agcctcgagg taaagagtca 900
tggaggtaga gcactgtttg gactaggggc ccttctcggg ttaccgaatt cagataaact 960
ccgaatgcca tgtacttata ctcgggagtc agactgcgag tgataagatc cgtagtcgaa 1020
agggaaacag cccagaccac cagttaaggt ccccaaatat atgttaagtg gaaaaggatg 1080
tggggttgct tagacaacca ggatgttggc ttagaagcag ccaccattga aagagtcgct 1140
aatagctcac tggtcgagtg accccgcgcc gaaaatgtac cggggctaaa catattaccg 1200
aaactgtgga tgaacctctt tagaggttcg tggtaggaga gcgttctaag ggcggtgaag 1260
tcagaccgga aggactggtg gagcgcttag aagtgagaat gccggtatga gtagcgaaag 1320
aagggtgaga atcccttcca ccgaatatct aaggtttcct gaggaaggct cgtccgctca 1380
gggttagtcg ggacctaagc cgaggccgat aggcgtaggc gatggacaac aggtagagat 1440
tcctgtacca gtgctaattg tttaaccgat ggggtgacac agaaggatag ggaatcgcac 1500
gaatggaaat gtgcgtccaa gcagtgagtg tgagaagtag gcaaatccgc ttctcacgaa 1560
gcatgagctg tgatggggaa ggaaattaag tacggaagtt cctgatttca cgctgtcaag 1620
aaaagcctct aggaagagta gtactgcccg taccgcaaac cgacacaggt agatgaggag 1680
agaatcctaa ggtgagcgag agaactctcg ttaaggaact cggcaaaatg accccgtaac 1740
ttcgggagaa ggggtgctct attagggtgc aagccccgaga gagccgcagt gaataggccc 1800
aggcgactgt ttagcaaaaa cacaggtctc tgcaaaaccg taaggtgacg tatagggggt 1860
gacgcctgcc cggtgctgga aggttaagag gagtgcttag cttcggcgaa ggtacgaatt 1920
gaagccccag taaacggcgg ccgtaactat aacggtccta aggtagcgaa attccttgtc 1980
gggtaagttc cgacccgcac gaaaggcgca acgatctggg cactgtctca acgagagact 2040
cggtgaaatt atagtacctg tgaagatgca ggttacccgc gacaggacgg aaagaccccg 2100
tggagcttta ctgcaacctg atatggaatg tttgtaccgc ttgtacagga taggtaggag 2160
ccgaagagac gtgtgcgcta gcatacgagg aggcaatggt gggatactac cctggctgta 2220
tgaccattct aacccgccac gcttagcgcg tggggagaca gtgtcaggtg ggcagtttga 2280
ctgggcggt cgcctcctaa agagtaacgg aggcgcccaa aggttccctc agaatggatg 2340
gaaatcattc gcagagtgta aaggcacaag ggagcttgac tgcgagactg acaagtcgag 2400
cagggacgaa agtcgggctt agtgatccgg tggttccgca tggaagggcc atcgctcaac 2460
ggataaaagc taccccgggg ataacaggct tatctccccc aagagtccac atcgacgggg 2520
aggtttggca cctcgatgtc ggctcgtcgc atcctggggc tgtagtcggt cccaagggtt 2580
gggctgttcg cccattaaag cggcacgcga gctgggttca gaacgtcgtg agacagttcg 2640
gtccctatcc gtcgcgggcg caggaaattt gagaggagct gtccttagta cgagaggacc 2700
gggatggaca caccgctggt gtaccagttg ttccgccagg agcatcgctg ggtagctatg 2760
tgtggcaggg ataaacgctg aaagcatcta agcgtgaagc ccccctcaag atgagatttc 2820
ccatttcttc ggaaagtaag atccctgaaa gatgatcagg tagataggtt tggagtggaa 2880
gtgtagcgat acatggagcg gacaaatact aatcgatcga ggacttaacc aa 2932

<210> 54
<211> 2923
<212> DNA
<213> Staphylococcus aureus

<400> 54
gattaagtta ttaagggcgc acggtggatg ccttggcact agaagccgat gaaggacgtt 60
actaacgacg atatgctttg gggagctgta agtaagcttt gatccagaga tttccgaatg 120
gggaaaccca gcatgagtta tgtcatgtta tcgatatgtg aatacatagc atatcagaag 180
gcacaccggg agaactgaaa catcttagta cccggaggaa gagaaagaaa attcgattcc 240
cttagtagcg gcgagcgaaa cgggaagagc ccaaaccaac aagcttgctt gttgggttg 300

taggacactc tatacggagt tacaaaggac gacattagac gaatcatctg gaaagatgaa 360
tcaaagaagg taataatcct gtagtcgaaa atgttgtctc tcttgagtgg atcctgagta 420
cgacggagca cgtgaaattc cgtcggaatc tgggaggacc atctcctaag gctaaatact 480
ctctagtgac cgatagtgaa ccagtaccgt gaggggaagg tgaaaagcac cccggaaggg 540
gagtgaaata gaacctgaaa ccgtgtgctt acaagtagtc agagcccgtt aatcgggtgat 600
ggcgtgcctt ttgtagaatg aaccggcgag ttacgatttg atgcaaggtt aagcagtaaa 660
tgtggagccg tagcgaaagc gagtctgaat agggcgttta gtatttggtc gtagacccga 720
aaccaggtga tctaccctg gtcaggttga agttccaggta acactgaatg gaggaccgaa 780
ccgacttacg ttgaaaagtg agcggatgaa ctgagggtag cggagaaatt ccaatcgaac 840
ctggagatag ctggttctct ccgaaatagc tttagggcta gcctcaagtg atgattattg 900
gaggtagagc actgtttgga cgaggggccc ctctcgggtt accgaattca gacaaactcc 960
gaatgccaat taatttaact tgggagtcag aacatgggtg ataaggtccg tgttcgaaag 1020
ggaaacagcc cagaccacca gctaaggtcc caaaatatat gttaagtgga aaaggatgtg 1080
gcgttgccca gacaactagg atgttggctt agaagcagcc atcatttaaa gagtgcgtaa 1140
tagctcacta gtcgagtgac actgcgccga aaatgtaccg gggctaaaca tattaccgaa 1200
gctgtggatt gtcctttgga caatggtagg agagcgttct aagggcgttg aagcatgatc 1260
gtaaggacat gtggagcgct tagaagtgag aatgccggtg tgagtagcga aagacgggtg 1320
agaatcccgt ccaccgattg actaaggttt ccagaggaag gctcgtccgc tctgggttag 1380
tcgggtccta agctgaggcc gacaggcgta ggcgatggat aacaggttga tattcctgta 1440
ccacctataa tcgttttaat cgatgggggg acgcagtagg ataggcgaag cgtgcgattg 1500
gattgcacgt ctaagcagta aggctgagta ttaggcaaat ccggtactcg ttaaggctga 1560
gctgtgatgg ggagaagaca ttgtgtcttc gagtcgttga tttcacactg ccgagaaaag 1620
cctctagata gaaaatagtt gcccgtacog caaaccgaca caggtagtca agatgagaat 1680
tctaaggtga gcgagcgaac tctcgttaag gaactcggca aaatgacccc gtaacttcgg 1740
gagaaggggt gctctttagg gttaacgccc agaagagccg cagtgaatag gcccaagcga 1800
ctgtttatca aaaacacagg tctctgctaa accgtaaggt gatgtatagg ggctgacgcc 1860
tgcccggtgc tggaaggtta agaggagtgg ttagcttctg cgaagctacg aatcgaagcc 1920
ccagtaaacg gcggccgtaa ctataacggt cctaaggtag cgaaattcct tgtcgggtaa 1980
gttccgaccc gcacgaaagg cgtaacgatt tgggcactgt ctcaacgaga gactcggtga 2040
aatcatagta cctgtgaaga tgcaggttac ccgcgacagg acggaaagac cccgtggagc 2100
tttactgtag cctgatattg aaattcggca cagcttgtac aggatagga ggagcctttg 2160
aaacgtgagc gctagcttac gtggaggcgc tggtgggata ctaccctagc tgtgttggct 2220
ttctaacccg caccacttat cgtggtggga gacagtgtca ggcgggcagt ttgactgggg 2280
cggtcgcctc ctaaaaggta acggaggcgc tcaaaggttc cctcagaatg gttggaaatc 2340
attcatagag tgtaaaggca taagggagct tgactgcgag acctacaagt cgagcagggt 2400
cgaaagacgg acttagtgat ccggtggttc cgcatggaag ggccatcgct caacggataa 2460
aagctacccc ggggataaca ggcttatctc ccccaagagt tcacatcgac ggggaggttt 2520
ggcacctcga tgtcggctca tcgcatcctg gggctgtagt cggtcccaag ggttgggctg 2580
ttcgcccatt aaagcggtac gcgagctggg ttcagaacgt cgtgagacag ttcggtccct 2640
atccgtcgtg ggcgtaggaa atttgagagg agctgtcctt agtacgagag gaccgggatg 2700
gacatacctc tggtgtacca gttgtcgtgc caacggcata gctgggtagc tatgtgtgga 2760
cgggataagt gctgaaagca tctaagcatg aagccccccct caagatgaga tttcccaact 2820
tcggttataa gatccctcaa agatgatgag gttaatagct tcgaggtgga agcatggtga 2880
catgtggagc tgacgaatac taatcgatcg aagacttaat caa 2923

<210> 55
<211> 2900
<212> DNA
<213> Streptococcus mutans

<400> 55
gttaagttaa taagggcgca cggtggatgc ctaggcacta ggagccgatg aaggacgtga 60
cgaacgacga catgcttttgg ggagctgtaa gtaagccttg atccagagat atccgaatgg 120
gggaacccaa caggtaatgc ctgttatcca taactgttaa ggttatgaga aggaagacgc 180
agtgaactga aacatctcag tagctgcagg aagagaaagc aagagcgatt gcctcagtag 240

cggcgagcga agaggcagga gggcaaacca gagtgtttac actctggggt tgtaggactg 300
cgataaagca gccaagggaa tagaagaaga ctctgggaag agtcgccaga gagagtaaga 360
gcctcgtatt tgaaattcac ttgatgccaa gcaggatcct gagtacggcg ggacacgagg 420
aatcccgtcg gaatctggga ggcccatctc ccaaccctaa atactcccta gtgaccgata 480
tgtgaaccagt accgtgaggg aaaggtgaaa agtaccccgg aaggtggagtg aaagagaacc 540
tgaaaccgtg tgcttacaag aagttcgagc ccgttaatgg gtgagagcgt gccttttgta 600
gaatgaaccg gcgagttacg tttacgtgcg aggttaagtt gaagagacgg agccgtagag 660
aaaccgagtc tgaaaagggc ggttaagtac gtagatgtag acccgaaacc aagtgaccta 720
cccatgagca ggttgaaggt gcggtaaaac gcactggagg accgaaccag gacacgttga 780
aaagtgtttg gatgacttgt gggtagcgga gaaattccaa acgaacttgg agatagctgg 840
ttctctccga aatagcttta gggctagcgt cggtcgcgag actcttggag gtagagcact 900
gtttgattga ggggtccatc ccggattacc aatctcagat aaactccgaa tgccaacgag 960
ttaagaccgg cagtcagact gcgagtgcta agatccgtag tcgaaaggga aacagcccag 1020
accaccagct aaggtcccca aataattgtt aagtggaaaa ggatgtgggg ttgcacagac 1080
aactaggatg ttagcttaga agcagctatt cattcaaaga gtgcgtaata gctcactagt 1140
cgagtgaccc tgcgccgaaa atgtaccggg gctgaaacaa tttaccgaag ctgtggatcc 1200
cttaggggat ggtaggagag cgttctatgt gcgcagaagg tgtaccgcaa ggagcgctgg 1260
agtgcataga agtgagaatg ccggtatgag tagcgtaaga caggtgagaa tcctgtccac 1320
cgtaagacta aggattccag gggaaggctc gtccgccctg ggttagtcgg gacctaagga 1380
gagaccgata ggtgtatccg atgggcaaca ggttgatatt cctgtactag agtattgagt 1440
gaaggaggga cgcagcaggc taactagagc gtgcgattgg aagagcacgt ccaagcagtg 1500
aggtgaggac tgagtcaaat gcttagttct gcgccaccaa gctgtgacgg ggagcgaagt 1560
ttagtagcga agctagtgat gtcactctgc caagaaaagc ttctagcgtt aatgaatact 1620
ctacccgtac cgcaaaccga cacaggtagt cgaggcgagt agcctcaggt gatcgagcga 1680
actctcgtta aggaactcgg caaaatggcc ccgtaacttc gggagaaggg gcgctggcga 1740
taagtcagcc gcagtgaaaa ggcccaagca actgtttatc aaaaacacag ctctctgcga 1800
aatcgtaaga tgaagtatag ggggtgacgc ctgcccggtg ctggaaggtt aagaggagcg 1860
cttagacgtt tgtcgaaggt gtgaattgaa gccccagtaa acggcggccg taactataac 1920
ggtcctaagg tagcgaaatt ccttgtcggg taagttccga cccgcacgaa aggcgtaatg 1980
atttgggcac tgtctcaacg agagactcgg tgaaatttta gtacctgtga agatgcaggt 2040
tacccgcgac aggacggaaa gaccccatgg agctttactg cagtttgata ttgcgtatct 2100
gttacacatg tacaggatag gtaggagcca aggaagagtg aacgctagtt tacttggagg 2160
cgttgttggg atactaccct tgtgtgatgg ctactctaac ccggtaggtt gatcatctac 2220
ggagacagtg tctgacgggc agtttgactg gggcggtcgc ctcctaaagc gtaacggagg 2280
cgcccaaagg ttccctcaga ctggttggaa atcagtcgta gagtgtaaag gtataaggga 2340
gcttgactgc gagacagaca agtcgagcag ggacgaaagt cgggcttagt gatccggtgg 2400
taccgtatgg aagggccatc gctcaacgga taaaagctac cctggggata acaggcttat 2460
ctcccccaag agttcacatc gacggggagg tttggcacct cgatgtcggc tcgtcgcatc 2520
ctggggctgt agtcggtccc aagggttggg ctgttcgccc attaaagcgg cacgcgagct 2580
gggttcagaa cgtcgtgaga cagttccggtc cctatccgtc gcgggcgaag gaaatttgag 2640
aggatctgct cctagtacga gaggaccaga gtggacttac cgctggtgta ccagttgttc 2700
tgccaagagc atcgctgggt agctaagtag ggaggggata aacgctgaaa gcatctaagt 2760
gtgaagcccc cctcaagatg agatttccca taacgttcag ttagtaagag ccctgaaaga 2820
agaacaggta gataggttgg gagtggaagc gttgtgagac gtgaagcgga ccaatactaa 2880
tcgctcgagg acttatccaa 2900

<210> 56
<211> 2902
<212> DNA
<213> Streptococcus pneumoniae

<400> 56
ggttaagtta ataagggcgc acggtggatg ccttggcact aggagccgac gaaggacgtg 60
acaaacgacg atatgccttg ggtagctgta agtaagcgat gatccaggga tttccgaatg 120
ggggaaccca acaggtaata cctgttaccc acatctgtta aggatgtgag gaggaagacg 180

cagtgaactg aaacatctaa gtagctgcag gaagagaaag caaaagcgat tgccttagta 240
gcggcgagcg aaacggcaga agggcaaacc gaagagttta ctcttcgggg ttgtaggact 300
gcaatgtgga ctcaaagatt atagaagaat gatttgggaa gatcagccaa agagagtaat 360
agcctcgtat ttaaaatagt ctttgtactt agcagtatcc tgagtacggc gggacacgtg 420
aaatcccgtc ggaatctggg aggaccatct cccaacccta aatactccct agtgacccgat 480
agtgaaccag taccgtgagg gaaaggtgaa aagcacccctg ggaggggagt gaaatagaac 540
ctgaaaccgt gtgcctacaa caagttcgag cccgttaatg ggtgagagcg tgccttttgt 600
agaatgaacc ggcgagttac gttatgatgc gaggttaagt tgaagagacg gagccgtagg 660
gaaaccgagt ctgaatagggg cgcctagta tcatgacgta gacccgaaac catgtgacct 720
acccatgagc aggttgaagg tgcggtaaga cgcactggag gaccgaacca gggcacgttg 780
aaaagtgctt ggatgacttg tgggtagcgg agaaattcca aacgaacttg gagatagctg 840
gttctctccg aaatagcttt agggctagcg tcgacattag agattcttgg aggtagagca 900
ctgtttgggt gaggggtcca tcccggatta ccaatctcag ataaactccg aatgccaatg 960
aattatggtc ggcagtcaga ctgcgagtgc taagatccgt agtcgaaagg gaaacagccc 1020
agaccaccag ctaaggtccc aaaataattg ttaagtggaa aaggatgtgg ggttgcacag 1080
acaactagga tgttagctta gaagcagcta ttcattcaaa gagtgcgtaa tagctcacta 1140
gtcgagtgac cctgcgccga aaatgtaccg gggctaaaac aatttaccga agctgtggat 1200
acctttatag gtatggtagg agagcgttct atgtgtgatg aaggtatacc gtgaggagtg 1260
ctggaacgca tagaagtgag aatgccggta tgagtagcga aagacaggtg agaatcctgt 1320
ccaccgtaag actaaggttt ccaggggaag gctcgtccgc cctgggttag tcgggaccta 1380
aggagagacc gaaaggtgta tccgatggac aacaggttga tattcctgta ctagagtatg 1440
tagtgatgga gggacgcagt aggctaacta aagcagacga ttggaagagt ctgtctaagc 1500
agtgaggtgt gaattgagtc aaatgcttaa ttctataaca ttgagctgtg atgggagcg 1560
aagtttagta gcgaagttag tgacgtcaca ctgccaagaa aagcttctag cgtttaaaca 1620
tactctaccc gtaccgcaaa ccgacacagg tagtcgaggc gagtagcctc aggtgagcga 1680
gagaactctc gttaaggaac tcggcaaaat gaccccgtaa cttcgggaga aggggtgctg 1740
acttaaagtc agccgcagtg aataggccca agcaactgtt tatcaaaaac acagctctct 1800
gctaaatcgt aagatgatgt atagggggtg acgcctgccc ggtgctggaa ggttaagagg 1860
agtgcttagc gtaagcgaag gtatgaattg aagccccagt aaacggcggc cgtaactata 1920
acggtcctaa ggtagcgaaa ttccttgtcg ggtaagttcc gacccgcacg aaaggcgtaa 1980
tgatttgggc actgtctcaa cgagagactc ggtgaaattt tagtacctgt gaagatgcag 2040
gttacccgcg acaggacgga aagaccccat ggagctttac tgcagtttga tattgagtgt 2100
ctgtaccaca tgtacaggat aggtaggagt ctaagagatc gggacgccag tttcgaagga 2160
gacgctgttg ggatactacc cttgtgttat ggccactcta acccagatag gtgatcccta 2220
tcggagacag tgtctgacgg gcagtttgac tggggcggtc gcctcctaaa aggtaacgga 2280
ggcgcccaaa ggttccctca gaatggttgg aaatcattcg cagagtgtaa aggtataagg 2340
gagcttgact gcgagagcta caactcgagc agggacgaaa gtcgggctta gtgatccggt 2400
ggttccgtat ggaagggcca tcgctcaacg gataaaagct accctgggga taacaggctt 2460
atctccccca agagttcaca tcgacgggga ggtttggcac ctcgatgtcg gctcgtcgca 2520
tcctggggct gtagtcggtc ccaaggttg ggctgttcgc ccattaaagc ggcacgcgag 2580
ctgggttcag aacgtcgtga gacagttcgg tccctatccg tcgcgggcgt aggaaatttg 2640
agaggatctg ctcctagtac gagaggacca gagtggactt accgctggtg taccagttgt 2700
cttgccaaag gcatcgctgg gtagctatgt agggaaggga taaacgctga aagcatctaa 2760
gtgtgaaacc cacctcaaga tgagatttcc catgattata tatcagtaag agccctgaga 2820
gatgatcagg tagataggtt agaagtggaa gtgtggcgac acatgtagcg gactaatact 2880
aatagctcga ggacttatcc aa 2902

<210> 57
<211> 2901
<212> DNA
<213> Streptococcus pyogenes

<400> 57
ggttaagtta ataagggcgc acggtggatg ccttggcact agaagccgaa gaaggacgtg 60
actaacgacg aaatgctttg gggagctgta agtaagcgct gatccagaga tgtccgaatg 120

ggggaacccg gcatgtaatg catgtcatcc atgactgtta aggtcatgag aaggaagacg 180
cagtgaactg aaacatctaa gtagctgcag gaagagaaag caaacgcgat tgccttagta 240
gcggcgagcg aaacggcagg agggcaaacc gaggagttta ctcctcgggg ttgtaggact 300
gcgaagtggg acataaagtt aatagaagaa ttacctggga aggtaagcca aagagagtaa 360
cagcctcgta tttaaaattg actttagccc tagcagtatc ctgagtacgg cgagacacgc 420
gaaatctcgt cggaatctgg gaggaccatc tccaaaccct aaatactctc tagtgaccga 480
tagtgaacca gtaccgtgag ggaaaggtga aaagcacccc gggaggggag tgaaatagaa 540
cctgaaaccg tgtgcctaca acaagttcga gcccgttaat gggtgagagc gtgccttttg 600
tagaatgaac cggcgagtta cgatatgatg cgaggttaag ttgaagagac ggagccgtag 660
ggaaaccgag tcttaatagg gcgtcatagt atcatgttgt agacccgaaa ccatgtgacc 720
taccccatgag caggttgaag gtgtggtaaa acgcactgga ggaccgaacc agggcacgtt 780
gaaaagtgct tggatgactt gtgggtagcg gagaaattcc aaacgaactt ggagatagct 840
ggttctctcc gaaatagctt tagggctagc gtcgatgtta agtctcttgg aggtagagca 900
ctgttgggt gaggggtcca tcccggatta ccaatctcag ataaactccg aatgccaacg 960
agatataatc ggcagtcaga ctgcgagtgc taagatccgt agtcgaaagg gaaacagccc 1020
agaccaccag ctaaggtccc aaaataactg ttaagtggaa aaggatgtgg ggttgcacag 1080
acaactagga tgttagctta gaagcagcta ttcattcaaa gagtgcgtaa tagctcacta 1140
gtcgagtgac cctgcgccga aaatgtaccg gggctaaaac agtttaccga agctgtggat 1200
gacacaaaag tgtcatggta ggagagcgtt ctatgtgtga agaaggtgta ccgtgaggag 1260
cgctggaacg catagaagtg agaatgccgg tatgagtagc gaaagacagg tgagaatcct 1320
gtccaccgta agactaaggt ttccagggga aggctcgtcc gccctgggtt agtcgggacc 1380
taaggagaga ccgaaaggtg tatccgatgg ccaacaggtt gatattcctg tactagagta 1440
tatagtgatg gagggacgca gtaggctaac taaaccggac gattggaaga gtccggctaa 1500
gcagtgaggt gtaagatgag tcaaatgctt atcttttaa cattgagctg tgatggggag 1560
cgaatttag tagcgaagtt agtgatgtca cactgccaag aaaagcttct agcgtttaat 1620
gatactctac ccgtaccgca aaccgacaca ggtagtcgag gcgagtagcc tcaggtgatc 1680
gagagaactc tcgttaagga actcggcaaa atgaccccgt aacttcggga gaaggggtgc 1740
tgacttaggt cagccgcagt gaataggccc aagcaactgt ttatcaaaaa cacagctctc 1800
tgctaaatcg taagatgatg tatagggggt gacgcctgcc cggtgctgga aggttaagag 1860
gagggtttag cgcaagcgaa gatctgaatt gaagccccag taaacggcgg ccgtaactat 1920
aacggtccta aggtagcgaa attccttgtc gggtaagttc cgacccgcac gaaaggcgta 1980
atgatttggg cactgtctca acgagagact cggtgaaatt ttagtacctg tgaagatgca 2040
ggttacccgc gacaggacgg aaagacccca tggagcttta ctgcagtttg atattgagta 2100
tctgtaccac atgtacagga taggtaggag ccattgactt cgggacgcca gtttcgaatg 2160
aggcgtcgtt gggatactac ccttgtgtta tggctactct aacccagata ggttatccct 2220
atcggagaca gtgtctgacg ggcagtttga ctggggcggt cgcctcctaa agagtaacgg 2280
aggcgcccaa aggttccctc agattggttg gaaatcaatc gcagagtgta aaggtataag 2340
ggagcttgac tgcgagagct acaactcgag cagggacgaa agtcgggctt agtgatccgg 2400
tggtaccgaa tggaagggcc atcgctcaac ggataaaagc taccctgggg ataacaggct 2460
tatctccccc aagagttcac atcgacgggg aggtttggca cctcgatgtc ggctcgtcgc 2520
atcctggggc tgtagtcggt cccaagggtt gggctgttcg cccattaaag cggcacgcga 2580
gctgggttca gaacgtcgtg agacagttcg gtccctatcc gtcgcgggcg taggaaattt 2640
gagaggatct gctcctagta cgaggaggacc agagtggact taccgctggt gtaccagttg 2700
tcttgccaaa ggcatcgctg ggtagctatg tagggaaggg ataagcgctg aaagcatcta 2760
agtgcgaagc ccccctcaag atgagatttc ccatgatttc atatcagtaa gagccctgag 2820
agatgatcag gtagataggt taggagtgta agtgtagcga tacatgtagc ggactaatac 2880
taatagctcg aggacttatc c 2901

<210> 58
<211> 3107
<212> DNA
<213> Mycobacterium avium

<400> 58
tgtgtgtaag taagtgttta agggcgcatg gtggatgcct tggcatcgag agccgatgaa 60

ggacgtggga ggctgcgata tgcctcgggg agctgtcaac cgagcattga tccgaggatt 120
tccgaatggg ggaacccagc acgagtgatg tcgtgttacc cgtatctgaa tatataggat 180
gcgggaggta acgcggggaa gtgaaacatc tcagtacccg taggagaaga aaacaattgt 240
gattccgtca gtagtggcga gcgaaccgga acaggctaaa ccgcatgcat ggacaaccgg 300
gtaggggttg tgtgtgcggg gttgtgggat tgatatgtct cagctctacc tggctgaggg 360
gtagtcagaa agtgtcgtgg ttagcggaag tggcctggga cggccccgcg tagaccggtga 420
gagcccggta cgcgaaaacc cggcacctgc cttatatcaa cacccgagta gcagcgggcc 480
cgtggaatct gctgtgaatc tgccgggacc acccggtaag cctaaatact tctcgatgac 540
cgatagcgga ttagtaccgt gagggaatgg tgaaaagtac cccgggaggg agtgaaatag 600
tacctgaaac cgtgtgccta caatccgtca gagcctcctc gtgggggtgat ggcgtgcctt 660
ttgaagaatg agcctgcgag tcagggacac gtcgcgaggt taaccgtgcc ggggtagccg 720
cagcgaaagc gagtctgaat agggcgcatc ccctttgggg tgtagtggcg tgttctggac 780
ccgaagcgga gtgatctacc catggccagg gtgaagcgcg ggtagagacg cgtggaggcc 840
cgaacccact taggttgaag actgagggga tgagctgtgg gtagggggtga aaggccaatc 900
aaatccggtg atagtcggtt ctccccgaaa tgcatttagg tgcagccgttg cgtggttcac 960
cacggaggta gagctactgg atggccgatg ggccctacta ggttactgac gtcagccaaa 1020
ctccgaatgc cgtggtgtaa aagcgtggca gtgagaccggc gggggataag ctccgtacgt 1080
cgaaagggaa acagcccaga tcgccggcta aggccccctaa gcgtgtgcta agtggaaaag 1140
gatgtgtagt cgcagagaca accaggaggt tggcttagaa gcagccatcc ttgaaagagt 1200
gcgtaatagc tcactggtca agtgattatg cgccgataat gtagcggggc tcaagcacac 1260
cgccgaagcc gcggcacatt catctttacg gtggatgtgg gtaggggagc gtccccccatt 1320
cagcgaagct ccgggtgacc ggtggtggag ggtgggggag tgagaatgca ggcatgagta 1380
gcgataaggc aagtgagaac cttgcccgcc gtaagaccaa gggttcctgg gccaggccag 1440
tccgcccagg gtgagtcggg acctaaggcg aggccgacag ggtagtcgat ggacaacggg 1500
ttgatattcc cgtacccgtg tatgggcgtc cctgatgaat cagcggtact aaccacccaa 1560
aaccggatcg accattcccc ttcgggggcg tggcgattcg gggctgcgtg ggaccttcgc 1620
tggtagtagt caagcaatgg ggtgacgcag gaaggcagcc gtaccagtca gtggtaatac 1680
tggggcaagc ccgtagagag cgatagggcaa atccgtcgct cactaatcct gagaggtgat 1740
gcatagccgg ttgaggcgaa ttcggtgatc ctctgctgcc aagaaaagcc tctagcgagc 1800
acatacacgg cccgtacccc aaaccaacac aggtggtcag gtagagaata ccaaggcgta 1860
cgagataact atggttaagg aactcggcaa aatgccccg taacttcggg agaaggggc 1920
ccggaatacc gtgaacaccc ttgcggtggg agcgggattc ggccgcagaa accagtgggt 1980
agcgactgtt tactaaaaac acaggtccgt gcgaagtcgc aagacgatgt atacggactg 2040
acgcctgccc ggtgctggaa ggttaagagg acccgttaac ccgtaagggt gaagcggaga 2100
atttaagccc cagtaaacgg cggtggtaac tataacccatc ctaaggtagc gaaattcctt 2160
gtcgggtaag ttccgacctg cacgaatggc gtaacgactt cccaactgtc tcaaccatag 2220
actcggcgaa attgcactac gagtaaagat gctcgttacg cgcggcagga cgaaaagacc 2280
ccgggacctt cactacaact tggtattggt gttcggtacg gtttgtgtag gatagggtgg 2340
agacttttgaa gcacagacgc cagtttgtgt ggagtcgttg ttgaaataacc actctgatcg 2400
tattggacac ctaacgtcga acccttatcg ggttcacgga cagtgcctgg cgggtagttt 2460
aactggggcg gttgcctcct aaaatgtaac ggaggcgccc aaaggttccc tcaacctgga 2520
cggcaatcag gtggcgagtg taagtgcaca agggagcttg actgcgagac ttacaagtca 2580
agcagggacg aaagtcggga ctagtgatcc ggcaccccg agtggaaggg gtgtcactca 2640
acggataaaa ggtaccccgg ggataacggg ctgatcttcc ccaagagtcc atatcgacgg 2700
gatggtttgg cacctcgatg tcggctcgtc gcatcctggg gctggagcag gtcccaaagg 2760
ttgggctgtt cgcccattaa agcggcacgc gagctgggtt tagaacgtcg tgagacagtt 2820
cggtctctat ccgccgcgcg cgtcagaaac ttgaggaaac ctgtccctag tacgagagga 2880
ccgggacgga cgaacctctg gtatcaccagt tgtcccacca ggggcacggc tggatagcca 2940
cgttcggaca ggataaccgc tgaaagcatc taagcgggaa accttctcca agatcaggtt 3000
tctcaccctt ttagagggat aaggcccccc gcagaccacg ggattgatag gccagacctg 3060
gaagctcagt aatgagtgca gggaactggc actaactggc cgaaagc 3107

<210> 59
<211> 3138
<212> DNA

<213> Mycobacterium tuberculosis

<400> 59

ttgtaagtgt ctaagggcgc atggtggatg ccttggcatc gagagccgat gaaggacgtg 60
ggagggtgcg atatgcctcg gggagctgtc aaccgagcgt ggatccgagg atttccgaat 120
ggggaaaccc agcacgagtg atgtcgtgct acccgcatct gaatatatag ggtgcgggag 180
ggaacgcggg gaagtgaaac atctcagtac ccgtaggagg agaaaacaat tgtgattccg 240
caagtagtgg cgagcgaacg cggaacaggc taaaccgcac gcatgggtaa ccgggtaggg 300
gttgtgtgtg cggggttgtg ggaggatatg tctcagcgct acccggctga gaggcagtca 360
gaaagtgtcg tggttagcgg aagtggcctg ggatggtctg ccgtagacgg tgagagcccg 420
gtacgcgaaa acccggcacc tgcctagtat caatttcccga gtagcagcgg gcccgtggaa 480
tccgctgtga atccgccggg accacccggt aagcctaaat actcctcgat gaccgatagc 540
ggattagtac cgtgagggaa tggtgaaaag tacccccgga ggggagtgaa agagtacctg 600
aaaccgtgtg cctacaatcc gtcagagcct ccttttcctc tccggaggag ggtggtgatg 660
gcgtgccttt tgaagaatga gcctgcgagt cagggacatg tcgcaaggtt aacccgtgtg 720
gggtagccgc agcgaaagcg agtctgaata gggcgaccca cacgcgcata cgcgcgtgtg 780
aatagtggcg tgttctggac ccgaagcgga gtgatctacc catggccagg gtgaagcgcg 840
ggtaagaccg cgtggaggcc cgaacccact taggttgaag actgaggggga tgagctgtgg 900
gtagggggtga aaggccaatc aaactccgtg atagctggtt ctccccgaaa tgcatttagg 960
tgcagcgttg cgtggttcac cgcggaggta gagctactgg atggccgatg ggccctacta 1020
ggttactgac gtcagccaaa ctccgaatgc cgtggtgtaa agcgtggcag tgagacggcg 1080
ggggataagc tccgtacgtc gaaagggaaa cagcccagat cgccggctaa ggccccccag 1140
cgtgtgctaa gtgggaaagg atgtgcagtc gcaaagacaa ccaggaggtt ggcttagaag 1200
cagccaccct tgaaagagtg cgtaatagct cactggtcaa gtgattgtgc gccgataatg 1260
tagcggggct caagcacacc gccgaagccg cggcacatcc accttgtggt gggtgtggat 1320
aggggagcgt ccctcattca gcgaagccac cgggtgaccg gtggtggagg gtggggggagt 1380
gagaatgcag gcatgagtag cgacaaggca agtgagaacc ttgcccgccg aaagaccaag 1440
ggttcctggg ccaggccagt ccgcccaggg tgagtcggga cctaaggcga ggccgacagg 1500
cgtagtcgat ggacaacggg ttgatattcc cgtacccgtg tgtgggcgcc cgtgacgaat 1560
cagcggtact aaccacccaa aaccggatcg atcactcccc ttcgggggtg tggagttctg 1620
gggctgcgtg ggaacttcgc tggtagtagt caagcgaagg ggtgacgcag gaaggtagcc 1680
gtaccagtca gtggtaacac tggggcaagc cggtagggag agcgataggc aaatccgtcg 1740
ctcactaatc ctgagaggtg acgcatagcc ggttgaggcg aattcggtga tcctctgctg 1800
ccaagaaaag cctctagcga gcacacacac ggcccgtacc ccaaaccgac acaggtggtc 1860
aggtagagca taccaaggcg tacgagataa ctatggttaa ggaactcggc aaaatgcccc 1920
cgtaacttcg ggagaagggg gaccggaata tcgtgaacac ccttgcggtg ggagcgggat 1980
ccggtcgcag aaaccagtga ggagcgactg tttactaaaa acacaggtcc gtgcgaagtc 2040
gcaagacgat gtatacggac tgacgcctgc ccggtgctgg aaggttaaga ggacccgtta 2100
acccgcaagg gtgaagcgga gaatttaagc cccagtaaac ggcggtggta actataacca 2160
tcctaaggta gcgaaattcc ttgtcgggta agttccgacc tgcacgaatg gcgtaacgac 2220
ttctcaactg tctcaaccat agactcggcg aaattgcact acgagtaaag atgctcgtta 2280
cgcgcggcag gacgaaaaga ccccgggacc ttcactacaa cttggtattg atgttcggta 2340
cggtttgtgt aggataggtg ggagactgtg aaacctcgac gccagttgag gcggagtcgt 2400
tgttgaaata ccactctgat cgtattgggc atctaacctc gaaccctgaa tcgggtttag 2460
ggacagtgcc tggcgggtag tttaactggg gcggttgcct cctaaaatgt aacggaggcg 2520
cccaaaggtt ccctcaacct ggacggcaat caggtggcga gtgtaaatgc acaaggggagc 2580
ttgactgcga gacttacaag tcaagcaggg acgaaagtcg ggattagtga tccggcacct 2640
ccgagtggaa ggggtgtcgc tcaacggata aaaggtaccc cggggataac aggctgatct 2700
tccccaagag tccatatcga cgggatggtt tggcacctcg atgtcggctc gtcgcatcct 2760
ggggctggag caggtcccaa gggttgggct gttcgcccat taaagcggca cgcgagctgg 2820
gtttagaacg tcgtgagaca gttcggtctc tatccgccgc gcgcgtcaga aacttggaga 2880
aacctgtccc tagtacgaga ggaccgggac ggacgaacct ctggtgcacc agttgtcccg 2940
ccaggggcac cgctggatag ccacgttcgg tcaggataac cgctgaaagc atctaagcgg 3000
gaaaccttct ccaagatcag gtttctcacc cacttggtgg gataaggccc cccgcagaac 3060
acgggttcaa taggtcagac ctggaagctc agtaatgggt gtagggaact ggtgctaacc 3120
ggccgaaaac ttacaaca 3138

<210> 60
<211> 2903
<212> DNA
<213> Escherichia coli

<400> 60
ggttaagcga ctaagcgtac acggtggatg ccctggcagt cagaggcgat gaaggacgtg 60
ctaatctgcg ataagcgtcg gtaaggtgat atgaaccgtt ataaccggcg atttccgaat 120
ggggaaaccc agtgtgattc gtcaactat cattaactga atccatagg taatgaggcg 180
aaccgggga actgaaacat ctaagtaccc cgaggaaaag aaatcaaccg agattccccc 240
agtagcggcg agcgaacggg gaggagccca gagcctgaat cagtgtgtgt gttagtggaa 300
gcgtctggaa aggcgcgcga tacagggtga cagccccgta cacaaaaatg cacatactgt 360
gagctcgatg agtagggcgg gacacgtggt atcctgtctg aatatgggag gaccatcctc 420
caaggctaaa tactcctgac tgaccgatag tgaaccagta ccgtgaggga aaggcgaaaa 480
gaaccccggc gaggggagtg aaaaagaacc tgaaaccgtg tacgtacaag cagtgggagc 540
ctcttttatg gggtgactgc gtacctttg tataatgggt cagcgactta tattctgtag 600
caaggttaac cgaatagggg agccgaaggg aaaccgagtc ttaaccgggc gttaagttgc 660
agggtataga cccgaaaccc ggtgatctag ccatgggcag gttgaaggtt gggtaacact 720
aactggagga ccgaaccgac taatgttgaa aaattagcgg atgacttgtg gctgggggtg 780
aaaggccaat caaaccggga gatagctggt tctccccgaa agctatttag gtagcgcctc 840
gtgaattcat ctccgggggt agagcactgt ttcggcaagg gggtcatccc gacttaccaa 900
cccgatgcaa actgcgaata ccggagaatg ttatcacggg agacatacgg cgggtgctaa 960
cgtccgtcgt gaagagggaa acaacccaga ccgccagcta aggtcccaaa gtcatggtta 1020
agtgggaaac gatgtgggaa ggcccagaca gccaggatgt tggcttagaa gcagccatca 1080
tttaaagaaa gcgtaatagc tcactggtcg agtcggcctg cgcggaagat gtaacggggc 1140
taaaccatgc accgaagctg cggcagcgac actgtgtgtt gttgggtagg ggagcgttct 1200
gtaagcctgt gaaggtgtac tgtgaggtat gctggaggta tcagaagtgc gaatgctgac 1260
ataagtaacg ataaagcggg tgaaaagccc gctcgccgga agaccaaggg ttcctgtcca 1320
acgttaatcg gggcagggtg agtcgacccc taaggcgagg ccgaaaggcg tagtcgatgg 1380
gaaacaggtt aatattcctg tacttggtgt tactgcgaag gggggacgga gaaggctatg 1440
ttggccgggc gacggttgtc ccggtttaag cgtgtaggct ggttttccag gcaaatccgg 1500
aaaatcaagg ctgaggcgtg atgacgaggc actacggtgc tgaagcaaca aatgccctgc 1560
ttccaggaaa agcctctaag catcaggtaa catcaaatcg taccccaaac cgacacaggt 1620
ggtcaggtag agaatacca ggcgcttgag agaactcggg tgaaggaact aggcaaaatg 1680
gtgccgtaac ttcgggagaa ggcacgctga tatgtaggtg aagtccctcg cggatggagc 1740
tgaaatcagt cgaagatacc agctggctgc aactgtttat taaaaacaca gcactgtgca 1800
aacacgaaag tggacgtata cggtgtgacg cctgcccggt gccggaaggt taattgatgg 1860
ggtcagcgca agcgaagctc ttgatcgaag ccccggtaaa cggcggccgt aactataacg 1920
gtcctaaggt agcgaaattc cttgtcgggt aagttccgac ctgcacgaat ggcgtaatga 1980
tggccaggct gtctcccacc gagactcagt gaaattgaac tcgctgtgaa gatgcagtgt 2040
acccgcggca agacggaaag accccgtgaa cctttactat agcttgacac tgaacattga 2100
gccttgatgt gtaggataagg tgggaggctt tgaagtgtgg acgccagtct gcatggagcc 2160
gaccttgaaa taccaccctt taatgtttga tgttctaacg tggaccccgtg atccggggttg 2220
cggacagtgt ctggtgggta gtttgactgg ggcggtctcc tcctaaagag taacggagga 2280
gcacgaaggt tggctaatcc tggtcggaca tcaggaggtt agtgcaatgg cataagccag 2340
cttgactgcg agcgtgacgg cgcgagcagg tgcgaaagca ggtcatagtg atccggtggt 2400
tctgaatgga agggccatcg ctcaacggat aaaaggtact ccggggataa caggctgata 2460
ccgcccaaga gttcatatcg acggcggtgt ttggcacctc gatgtcggct catcacatcc 2520
tggggctgaa gtaggtccca agggtatggc tgttcgccat ttaaagtggt acgcgagctg 2580
ggtttagaac gtcgtgagac agttcggtcc ctatctgccg tgggcgctgg agaactgagg 2640
gggggctgctc ctagtacgag aggaccggag tggacgcatc actggtgttc gggttgtcat 2700
gccaatggca ctgcccggta gctaaatgcg gaagagataa gtgctgaaag catctaagca 2760
cgaaacttgc cccgagatga gttcttccctg accctttaag ggtcctgaag gaacgttgaa 2820
gacgacgacg ttgataggcc gggtgtgtaa gcgcagcgat gcgttgagct aaccggtact 2880

aatgaaccgt gaggcttaac ctt 2903

<210> 61
<211> 2903
<212> DNA
<213> Klebsiella pneumoniae

<400> 61
ggttaagcga ctaagcgtac acggtggatg ccctggcagt cagaggcgat gaaggacgtg 60
ctaatctgcg aaaagcgtcg gtaaggtgat atgaaccgtt ataaccggcg atgtccggat 120
ggggaaaccc agtgcaattc gttgcactat cgttaactga atacataggt taacgaggcg 180
aaccggggga actgaaacat ctaagtaccc cgaggaaaag aaatcaaccg agattcccac 240
agtagcggcg agcgaacggg gagcagccca gagtctgaat cagcttgtgt gttagtggaa 300
cggtctggaa agtccgacgg tcagaggtga tagtcccgta caccaaaatg cacaggctgt 360
gaactcgaag agtagggcgg gacacgtggt atcctgtctg aatatggggg gaccatcctc 420
caaggctaaa tactcctgac tgaccgatag tgaaccagta ccgtgagggа aaggcgaaaa 480
gaaccccggc gaggggagtg aaaaagaacc tgaaaccgtg tacgtacaag cagtgggagc 540
accttcgggt gtgactgcgt acctttтgta taatgggtca gcgacttata ttctgtagca 600
aggttaaccg tataggggag ccgcagggaa accgagtctt aactgggcgt taagttgcag 660
ggtatagacc cgaaacccgg tgatctagcc atgggcaggt tgaaggttgg gtaacactaa 720
ctggaggacc gaaccgacta atgttgaaaa attagcggat gacttgtggc tgggggtgaa 780
aggccaatca aaccgggaga tagctggttc tccccgaaag ctatttaggt agcgcctcgt 840
gaactcatct tcgggggtag agcactgttt cggctagggg gtcatcccga cttaccaacc 900
cgatgcaaac tacgaatacc gaagaatgtt atcacgggag acacacggcg ggtgctaacg 960
tccgtcgtga agagggaaac aacccagacc gccagctaag gtcccaaagt catggttaag 1020
tgggaaacga tgtgggaagg cacagacagc caggatgttg gcttagaagc agccatcatt 1080
taaagaaagc gtaatagctc actggtcgag tcggcctgcg cggaagatgt aacggggcta 1140
aaccatgcac cgaagctgcg gcagcgacac tatgtgttgt tgggtagggg agcgtttctgt 1200
aagcctgcga aggtgtgctg tgaggcatgc tggaggtatc agaagtgcga atgctgacat 1260
aagtaacgat aaagcgggtg aaaagcccgc tcgccggaag accaagggtt cctgtccaac 1320
gttaatcggg gcagggtgag tcgaccccta aggcgaggcc gaaaggcgta gtcgatggga 1380
aacaggttaa tattcctgta cttggtgtta ctgcgaaggg gggacggaga aggctatgtt 1440
agccgggcga cggttgtccc ggtttaagca tgtaggctgg ttgtccaggc aaatccggat 1500
aatcaaggct gaggtgtgat gacgaggcac tacggtgctg aagtaacaaa tgctctgctt 1560
ccaggaaaag cctctaagca tcaggtaaca tcaaatcgta ccccaaaccg acacaggtgg 1620
tcaggtagag aatacccaagg cgcttgagat aactcgggtg aaggaactag gcaaaatggt 1680
gccgtaactt cgggagaagg cacgctggtg tgtaggtgaa gcccctgccg ggtggagctg 1740
agaccagtcg aagataccag ctggctgcaa ctgtttatta aaaacacagc actgtgcaaa 1800
cacgaaagtg gacgtatacg gtgtgacgcc tgcccggtgc cggaaggtta attgatgggg 1860
ttatccgtaa ggagaagctc ttgatcgaag ccccggtaaa cggcggccgt aactataacg 1920
gtcctaaggt agcgaaattc cttgtcgggt aagttccgac ctgcacgaat ggcgtaatga 1980
tggccaggct gtctccaccc gagactcagt gaaattgaac tcgctgtgaa gatgcagtgt 2040
acccgcggca agacggaaag accccgtgaa cctttactat agcttgacac tgaacattga 2100
gccttgatgt gtaggatagg tgggaggctt tgaagcgtgg acgccagtct gcgtggagcc 2160
aaccttgaaa taccaccctt taatgtttga tgttctaacg ttggcccctc accgggggttg 2220
cggacagtgt ctggtgggta gtttgactgg ggcggtctcc tcccaaagcg taacggagga 2280
gcacgaaggt tagctaatcc tggtcggaca tcaggaggtt agtgcaatgg cataagctag 2340
cttgactgcg agcgtgacgg cgcgagcagg tgcgaaagca ggtcatagtg atccggtggt 2400
tctgaatgga agggccatcg ctcaacggat aaaaggtact ccggggataa caggctgata 2460
ccgccaaga gttcatatcg acggcggtgt ttggcacctc gatgttcggct catcacatcc 2520
tggggctgaa gtaggtccca agggtatggc tgttcgccat ttaaagtggt acgcgagctg 2580
ggttagaac gtcgtgagac agttcggtcc ctatctgccg tgggcgctgg agaattgagg 2640
ggggctgctc ctagtacgag aggaccggag tggacgcatc actggtgttc gggttgtcat 2700
gccaatggca ctgcccggta gctaaatgcg gaagagataa gtgctgaaag catctaagca 2760
cgaaacttgc cccgagatga gttcttccctg agacttttaag tctcctgaag gaacgttgaa 2820

gacgacgacg ttgataggcc gggtgtgtaa gcgcagcgat gcgttgagct aaccggtact 2880
aatgaaccgt gaggcttaac ctt 2903

<210> 62
<211> 2897
<212> DNA
<213> Haemophilus influenzae

<400> 62
gtatagttaa gtgactaagc gtacaaggtg gatgccttgg caatcagagg cgaagaagga 60
cgtgctaatc tgcgaaaagc ttggatgagt cgataagagg cgtttaatcc aagatatccg 120
aatggggaaa cccagtagat gaagaatcta ctatcaacaa gtgaattcat agcttgttga 180
ggcaaaccgg gagaactgaa acatctaagt accccgagga aaagaaatca accgagattt 240
cgtcagtagc ggcgagcgaa agcgaaagag ccagtaagtg atagcaatat agtgaggaga 300
atgtgttggg aagcacaatc aaagagggtg ataatcccgt atctaaaaac catattgtgg 360
tactaagcta acgagaagta gggcgggaca cgtgatatcc tgtttgaaga agggggccc 420
atcctccaag gctaaatact cctgattgac cgatagtgaa ccagtactgt gaaggaaagg 480
cgaaaagaac cccggtgagg ggagtgaaat agaacctgaa accttgtacg tacaagcagt 540
gggagcgagg gcaaccttgt gactgcgtac cttttgtata atgggtcagc gacttatatt 600
ttgtagcgag gttaaccgaa taggggagcc gaagggaaac cgagtcttaa ctgggcgaat 660
agttgcaagg tatagacccg aaacccggtg atctagccat gggcaggttg aaggttgggt 720
aacactaact ggaggaccga accgactaat gttgaaaaat tagcggatga cttgtggctg 780
ggggtgaaag gccaatcaaa ccgggagata gctggttctc cccgaaatct atttaggtag 840
agccttgagg tgacaccttt gggggtagag cactgtttcg gctagggggc catcccggct 900
taccaacccg atgcaaacta cgaataccaa agagtgatac tcaggagaca cacggcgggt 960
gctaacgtcc gtcgtggaga gggaaacaac ccagaccgcc agctaaggtc cccaagtcta 1020
tattaagtgg gaaacgaagt gggaaggctt agacagctag gatgttggct tagaagcagc 1080
catcatttaa agaaagcgta atagctcact agtcgagtcg gcctgcgcgg aagatgtaac 1140
gggggctgaaa tatagcaccg aagctgcggc atcagaattt attctgttgg gtaggggagc 1200
gttgtgtaag cggaagaagg ttcatcgaga ggtgggctgg acgtatcaca agtgcgaatg 1260
ctgacataag taacgataaa acgggtgaaa aacccgttcg ccggaagacc aagggttcct 1320
gtccaacgtt aatcggggca gggtgagtcg gctcctaagg cgaggctgaa aagcgtagtc 1380
gatgggaaac aggttaatat tcctgtactt ggtaaagctg cgatgtgggg acggagtagg 1440
ttaggttatc gcactgttgg atatgtgcgt ttaagttggt aggtgggaag tttaggcaaa 1500
tccggacttc cttaacacag agagatgatg acgaggctct acggagctga agtaactgat 1560
accacacttc caggaaaagc cactaagcga aaggctttac taaaccgtac tgaaaaccga 1620
cacaggtggt caggtagaga atactcaggc gcttgagaga actcgggtga aggaactagg 1680
caaaatagca ccgtaacttc gggagaaggt gcgccggcgt agattgtaag ggctagcccc 1740
tgaaggttga accggtcgaa gataccacgct ggctgcaact gtttattaaa aacacagcac 1800
tctgcaaaca cgaaagtgga cgtatagggt gtgatgcctg cccggtgctg gaaggttaat 1860
tgatggtgtc atcgaaagag aagcacctga tcgaagcccc agtaaacggc ggccgtaact 1920
ataacggtcc taaggtagcg aaattccttg tcgggtaagt tccgacctgc acgaatggca 1980
taatgatggc caggctgtct ccacccgaga ctcagtgaaa ttgaaatcgc cgtgaagatg 2040
cggtgtaccc gcggctagac ggaaagaccc cgtgaacctt tactatagct tgacactgaa 2100
cattgaattt tgatgtgtag gataggtggg agccttttgaa gcagtgacgc cagtcattgt 2160
ggaggcgacc ttgaaatacc acccttaac gtttgatgtt ctaacgaaga tgacgaaacg 2220
tggtctcgga cagtgtctgg tgggtagttt gactggggcg gtctcctccc aaagcgtaac 2280
ggaggagcac gaaggtttgc taatcacggt cggacatcgt gaggttagtg caatggtata 2340
agcaagctta actgcgagac agacaagtcg agcaggtacg aaagtaggtc atagtgatcc 2400
ggtggttctg aatggaaggg ccatcgctca acggataaaa ggtactccgg ggataacagg 2460
ctgataccgc ccaagagttc atatcgacgg cggtgtttgg cacctcgatg tcggctcatc 2520
acatcctggg gctgaagtag gtcccaaggg tatggctgtt cgccatttaa agtggtacgc 2580
gagctgggtt tagaacgtcg tgagacagtt cggtccctat ctgccgtggg cgtaggatga 2640
ttgattgggg ctgctcctag tacgagagga ccggagtgga cgcatcactg gtgttccggt 2700
tgtgtcgcca gacgcattgc cgggtagcta aatgcggaag agataagtgc tgaaagcatc 2760

taagcacgaa acttgccaag agatgagtca tccctgactt taagtcagta aggattgttg 2820
tagactacga cgtagatagg ttgggtgtgt aagtgatgtg agtcattgag ctaaccaata 2880
ctaattgccc gagaggc 2897

<210> 63
<211> 2865
<212> DNA
<213> Bordetella bronchiseptica

<220>
<221> modified_base
<222> (622)
<223> N = A, C, G or T/U

<400> 63
gatcaagcga ctaagtgcat atggtggatg ccttggcgat cacaggcgga tgaaggacgt 60
agtagcctgc gaaaagctgc ggggagctgg caaacaagca ttgatccgca gatatccgaa 120
tggggaaacc cacggcaagc ggtatccctg gctgaataca taggccagtg gaggcgaacc 180
gggtgaactg aaacatctca gtagctcgag gaaaagaaat caaccgagat tccgaaagta 240
gtggcgagcg aaatcggaag agcctttacg atttagcatt ttgcatagtc gaacggaatg 300
gaaagtccgg ccgtagcagg tgatagccct gtagacgaat gcagagtgtg gaactaggcg 360
taagagaagt agggcgggac acgtgaaatc ctgtctgaag atgggggtac catcctccaa 420
ggctaaatac tcgtgatcga ccgatagtga accagtaccg tgaggaaagg cgaaaagaac 480
cccggaagga gtgaaatagа tcctgaaacc gtatgcatac aacagtcgga gcctctttat 540
ggggtgacgg cgtacctttt gtataatggg tcagcgactt acattcagtg gcagcttaac 600
cgaatagga aggcgtcaga anagcagtcc gaataggcg ttccagtcgc tgggtgtaga 660
cccgaaacca gatgatctac ccatggccag gttgaaggca cggtaacacg tgctggagga 720
ccgaacccac tagtgttgaa aaactagggg atgagctgtg gataggggtg aaaggctaaa 780
caaatctgga aatagctggt tctctccgaa aactatttag gtagtgcctc aagtattact 840
gcaggggtta gagcactgtt atggctaggg ggtcatggcg acttaccaaa ccatggcaaa 900
ctccgaatac ctgcaagtac agcttgggag acagacgacc gggtgctaac gtccggactc 960
aagagggaaa caacccagac cgccagctaa ggtcccgaat tatcgctaag tgggaaacga 1020
agtgggaagg catagacagt caggaggttg gcttagaagc agccaccctt taaagaaagc 1080
gtaatagctc actgatcgag tcgtcctgcg cggaagatgt aacggctaag cgataaaccg 1140
aagctgcggg tgtgcacttt tagtgcagcg gtaggagagc gttctgtaag cctgcgaagg 1200
tggcttgtaa aggctgctgg aggtatcaga agtgcgaatg ctgacatgag tagccatаaa 1260
gggggtgaaa agccccctcg ccgtaagtcc aaggtttcct gcgcaacgtt catcggcgca 1320
gggtgagtcg gcccctaagg cgaggcagag atgcgtagct gatgggaagc tggttaatat 1380
tccagcaccg tcgtacagtg cgatgggggg acggatcgcg gaaggtcatc agggtgttgg 1440
acgtccctgt tgctgcattg aagatggcgc ttaggcaaat ccgggcgcga gaatcaaggg 1500
tgtggcacga gcgagcaagt ctcgcgaagt gattggaagt ggttccaaga aaagcctcta 1560
agcttcagct gtacgagacc gtaccgcaaa ccgacacagg tgggacggga tgaatattcc 1620
aaggcgcttg agagaactca ggagaaggaa ctcggcaaat tgataccgta acttcgggag 1680
aaggtatacc ctggtagtgt gaagcctgcg cgctgagcat gaaggggtcg cagagaatcg 1740
gtggctgcga ctgtttatta aaaacacagc actctgcaaa gacgaaagtc gacgtatagg 1800
gtgtgacgcc tgcccggtgc cggaaggtta agtgatgggg tgcaagctct tgatcgaagc 1860
cccggtaaac ggcggccgta actataacgg tcctaaggta gcgaaattcc ttgtcgggta 1920
agttccgacc tgcacgaatg gcgtaacgat ggccacactg tctcctcctg agactcagcg 1980
aagttgaagt gtttgtgatg atgcaatcta cccgcggcta gacggaaaga ccccatgaac 2040
ctttactgta gctttgcatt ggactgtgaa ccggcctgtg taggataggt gggaggcgca 2100
gaactcgagt cgccagattc gagggagcca tcctgaaat accaccctgg tttgttgcg 2160
gttctaacct tggtccgtta tccggatcgg ggacagtgca tggtaggcag tttgactggg 2220
gcggtctcct cccaaagcgt aacggaggag ttcgaaggta cgctaggtac ggtcggaaat 2280
cgtgctgata gtgcaatggc ataagcgtgc ttgactgtga gactgacagt gaacaggtgc 2340
gaacgggaca tagtgatccg gtggttctga tggaagggcc atcgctcaac ggataaaggt 2400

```
actctgggat aacaggctga taccgcccaa gagttcatat cgacggcggt gtttggcacc 2460
tcgatgtcgg ctcatctcat cctggggctg tagccggtcc aagggtatgc tgttcgccat 2520
ttaaagaggt acgtgagctg ggttagaaa cgtcgtgaga cagttttggtc cctatctgcc 2580
gtgggcgttg gatacttgaa caggagcctg ctcctagtac gagaggacgg gagtggacgt 2640
acctctggtg taccggttgt catgccaatg gcattgccgg gtagctaagt acggaagaga 2700
taaccgctga aggcatctaa gcggggaaact cgtctgaaga ttaggtatcc cggggactag 2760
atccccctga agggtcgttc gagaccagga cgttgatagg tcgggtgtgg aagcgcagta 2820
atgcgttaag ctaaccgata ctaattgccc gtgaggctta atcct                  2865

<210> 64
<211> 2865
<212> DNA
<213> Bordetella parapertussis

<220>
<221> modified_base
<222> (624)
<223> N = A, C, G or T/U

<400> 64
gatcaagcga ctaagtgcat atggtggatg ccttggcgat cacaggcgat gaaggacgta 60
gtagcctgcg aaaagctgcg gggagctggc aaacaagcat tgatccgcag atatccgaat 120
ggggaaaccc acggcaagcg gtatccctgg ctgaatacat aggccagtgg aggcgaaccg 180
ggtgaactga aacatctcag tagctcgagg aaaagaaatc aaccgagatt ccgaaagtag 240
tggcgagcga aatcggaaga gcctttacga tttagcattt tgcatagtcg aacggaatgg 300
aaagtccggc cgtagcaggt gatagccctg tagacgaaat gcagagtgtg gaactaggcg 360
taagagaagt agggcgggac acgtgaaatc ctgtctgaag atgggggggac catcctccaa 420
ggctaaatac tcgtgatcga ccgatagtga accagtaccg tgaggaaagg cgaaaagaac 480
cccggaagga gtgaaataga tcctgaaacc gtatgcatac aaacagtcgg agcctcttta 540
tggggtgacg gcgtaccttt tgtataatgg gtcagcgact tacattcagt ggcgagctta 600
accgaatagg gaaggcgtca gaanagcagt ccgaataggg cgtccagtcg ctgggtgtag 660
acccgaaacc agatgatcta cccatggcca ggttgaaggc acggtaacac gtcgtggagg 720
accgaaccca ctagtgttga aaaactaggg gatgagctgt ggataggggt gaaaggctaa 780
acaaatctgg aaatagctgg ttctctccga aaactattta ggtagtgcct caagtattac 840
tgcagggggt agagcactgt tatggctagg gggtcatggc gacttaccaa accatggcaa 900
actccgaata cctgcaagta cagcttggga gacagacgac cgggtgctaa cgtccggact 960
caagagggaa acaacccaga ccgccagcta aggtcccgaa ttatcgctaa gtgggaaacg 1020
aagtgggaag gcatagacag tcaggaggtt ggcttagaag cagccaccct ttaaagaaag 1080
cgtaatagct cactgatcga gtcgtcctgc gcggaagatg taacggctaa gcgatcaaacc 1140
gaagctgcgg gtgtgcactt ttagtgcagc ggtaggagag cgttctgtaa gcctgcgaag 1200
gtggcttgta aaggctgctg gaggtatcag aagtgcgaat gctgacatga gtagccgataa 1260
aggggtgaa aagccccctc gccgtaagtc caaggtttcc tgcgcaacgt tcatcggcgc 1320
agggtgagtc ggcccctaag gcgaggcaga gatgcgtagc tgatgggaag ctggttaata 1380
ttccagcacc gtcgtacagt gcgatggggg gacggatcgc ggaaggtcat cagggtgttg 1440
gacgtccctg ttgctgcatt gaagatggcg cttaggcaaa tccgggcgcg agaatcaagg 1500
gtgtggcacg agcgagcaag tctcgcgaag tgattggaag tggttccaag aaaagcctct 1560
aagcttcagc tgtacgagac cgtaccgcaa accgacacag gtgggacggg atgaatattc 1620
caaggcgctt gagagaactc aggagaagga actcggcaaa ttgatacccgt aacttcggga 1680
gaaggtatac cctggtagtg tgaagcctgc gcgctgagca tgaaggggtc gcagagaatc 1740
ggtggctgcg actgtttatt aaaaacacag cactctgcaa agacgaaagt cgacgtatag 1800
ggtgtgacgc ctgcccggtg ccggaaggtt aagtgatggg gtgcaagctc ttgatcgaag 1860
ccccggtaaa cggcggccgt aactataacg gtcctaaggt agcgaaattc cttgtcgggt 1920
aagttccgac ctgcacgaat ggcgtaacga tggccacact gtctcctcct gagactcagc 1980
gaagttgaag tgtttgtgat gatgcaatct acccgcggct agacggaaag accccatgaa 2040
cctttactgt agctttgcat tggactgtga accggcctgt gtaggatagg tgggaggcgc 2100
```

```
agaactcgag tcgccagatt cgagggagcc atccttgaaa taccaccctg gtttgtttgc 2160
ggttctaacc ttggtccgtt atccggatcg gggacagtgc atggtaggca gtttgactgg 2220
ggcggtctcc tcccaaagcg taacggagga gttcgaaggt acgctaggta cggtcggaaa 2280
tcgtgctgat agtgcaatgg cataagcgtg cttgactgtg agactgacag tcgaacaggt 2340
gcgaacggga catagtgatc cggtggttct gatggaaggg ccatcgctca acggataaag 2400
gtactctggg ataacaggct gataccgccc aagagttcat atcgacggcg gtgtttggca 2460
cctcgatgtc ggctcatctc atcctggggc tgtagccggt ccaaggggtat gctgttcgcc 2520
atttaaagag gtacgtgagc tgggtttaga aacgtcgtga gacagtttgg tccctatctg 2580
ccgtgggcgt tggatacttg aacaggagcc tgctcctagt acgagaggac cggagtggac 2640
gtacctctgg tgtaccggtt gtcatgccaa tggcattgcc gggtagctaa gtacggaaga 2700
gataaccgct gaaggcatct aagcggaaac tcgtctgaag attaggtatc ccgggactag 2760
atccccctga agggtcgttc gagaccagga cgttgatagg tcgggtgtgg aagcgcagta 2820
atgcgttaag ctaaccgata ctaattgccc gtgaggcttg atcct                  2865
```

<210> 65
<211> 2864
<212> DNA
<213> Bordetella pertussis

<220>
<221> modified_base
<222> (624)
<223> N = A, C, G or T/U

<400> 65

```
gatcaagcga ctaagtgcat atggtggatg ccttggcgat cacaggcgat gaaggacgta 60
gtagcctgcg aaaagctgcg gggagctggc aaacaagcat tgatccgcag atatccgaat 120
ggggaaaccc acggcaagcg gtatccctgg ctgaatacat aggccagtgg aggcgaaccg 180
ggtgaactga aacatctcag tagctcgagg aaaagaaatc aaccgagatt ccgaaagtag 240
tggcgagcga aatcggaaga gcctttacga tttagcattt tgcatagtcg aacggaatgg 300
aaagtccggc cgtagcaggt gatagccctg tagacgaaat gcagagtgtg gaactaggcg 360
taagagaagt agggcgggac acgtgaaatc ctgtctgaag atgggggggac catcctccaa 420
ggctaaatac tcgtgatcga ccgatagtga accagtaccg tgaggaaagg cgaaaagaac 480
cccggaagga gtgaaataga tcctgaaacc gtatgcatac aaacagtcgg agcctcttta 540
tggggtgacg gcgtaccttt tgtataatgg gtcagcgact tacattcagt ggcgagctta 600
accgaatagg gaaggcgtca gaanagcagt ccgaataggg cgtccagtcg ctgggtgtag 660
acccgaaacc agatgatcta cccatggcca ggttgaaggc acggtaacac gtcgtggagg 720
accgaaccca ctagtgttga aaaactaggg gatgagctgt ggataggggt gaaaggctaa 780
acaaatctgg aaatagctgg ttctctccga aaactattta ggtagtgcct caagtattac 840
tgcagggggt agagcactgt tatggctagg gggtcatggc gacttaccaa accatggcaa 900
actccgaata cctgcaagta cagcttggga gacagacgac cgggtgctaa cgtccggact 960
caagaggggaa acaacccaga ccgccagcta aggtcccgaa ttatcgctaa gtgggaaacg 1020
aagtgggaag gcatagacag tcaggaggtt ggcttagaag cagccaccct ttaaagaaag 1080
cgtaatagct cactgatcga gtcgtcctgc gcggaagatg taacggctaa gcgataaacc 1140
gaagctgcgg gtgtgcactt ttagtgcagc ggtaggagag cgttctgtaa gcctgcgaag 1200
gtggcttgta aaggctgctg gaggtatcag aagtgcgaat gctgacatga gtagcgataa 1260
agggggtgaa aagccccctc gccgtaagtc caaggtttcc tgcgcaacgt tcatcggcgc 1320
agggtgagtc ggcccctaag gcgaggcaga gatgcgtagc tgatgggaag ctggttaata 1380
ttccagcacc gtcgtacagt gcgatggggg gacggatcgc ggaaggtcat cagggtgttg 1440
gacgtccctg ttgctgcatt gaagatggcg cttaggcaaa tccgggcgcg agaatcaagg 1500
gtgtggcacg agcgagcaag tctcgcgaag tgattggaag tggttccaag aaaagcctct 1560
aagcttcagc tgtacgagac cgtaccgcaa accgacacag gtgggacggg atgaatattc 1620
caaggcgctt gagagaactc aggagaagga actcggcaaa ttgatacccgt aacttcggga 1680
gaaggtatac cctggtagtg tgaagcctgc gcgctgagca tgaaggggtc gcagagaatc 1740
ggtggctgcg actgtttatt aaaaacacag cactctgcaa agacgaaagt cgacgtatag 1800
```

ggtgtgacgc ctgcccggtg ccggaaggtt aagtgatggg gtgcaagctc ttgatcgaag 1860
ccccggtaaa cggcggccgt aactataacg gtcctaaggt agcgaaattc cttgtcgggt 1920
aagttccgac ctgcacgaat ggcgtaacga tggccacact gtcctcctct gagactcagc 1980
gaagtgaag tgttgtgtat gatgcaatct acccgcggct agacggaaag accccatgaa 2040
ccttttactgt agcttgcgat tggactgtga accggcctgt gtaggatagg tggaggcgc 2100
agaactcgag tcgccagatt cgagggagcc atcgttgaaa taccaccctg gtttgtttgc 2160
ggttctaacc ttgttccgtt atccggatcg gggacagtgc atggtaggca gtttgactgg 2220
ggcggtctcc tcccaaagcg taacggagga gttcgaaggt acgctaggta cggtcggaaa 2280
tcgtgctgat agtgcaatgg cataagcgtg cttgactgtg agactgacag tcgaacaggt 2340
gcgaacggga catagtgatc cggtggttct gatggaaggg ccatcgctca acggataaag 2400
gtactctggg ataacaggct gataccgccc aagagttcat atcgacggcg gtgtttggca 2460
cctcgatgtc ggctcatctc atcctggggc tgtagccggt ccaagggtat gctgttcgcc 2520
atttaaagag gtacgtgagc tgggtttaaa acgtcgtgag acagtttggt ccctatctgc 2580
cgtgggcgtt ggatacttga acaggagcct gctcctagta cgagaggacc ggagtggacg 2640
tacctctggt gtaccggttg tcatgccaat ggcattgccg ggtagctaag tacggaagag 2700
ataaccgctg aaggcatcta agcggaaact cgtctgaaga ttaggtatcc cgggactaga 2760
tcccccctgaa gggtcgttcg agaccaggac gttgataggt cgggtgtgga agcgcagtaa 2820
tgcgttaagc taaccgatac taattgcccg tgaggcttga tcct 2864

<210> 66
<211> 2878
<212> DNA
<213> Burkholderia cepacia

<400> 66
ggtcaagcga acaagtgcat gtggtggatg ccttggcgat cacaggcgat gaaggacgcg 60
gtagcctgcg aaaagctacg gggagctggc aaacaagctt tgatccgtag atgtccgaat 120
ggggaaaccc actccttttg gagtatccat ggctgaatac ataggccatg cgaaggaacg 180
cggtgaactg aaacatctaa gtaaccgcag gaaaagaaat caaccgagat tcccaaagta 240
gtggcgagcg aaatgggatg agccttgcac tctttatttg tattgttagc cgaacgctct 300
ggaaagtgcg gccatagcag gtgatagccc tgtaggcgaa aacagtatga aagaactagg 360
tgtgcgacaa gtagggcggg acacgtgaaa tcctgtctga agatgggggg accatcctcc 420
aaggctaaat actcgtgatc gaccgatagt gaaccagtac cgtgagggaa aggcgaaaag 480
aaccccggga ggggagtgaa atagatcctg aaaccgcatg catacaaaca gtcggagcct 540
cgtaaggggt gacggcgtac cttttgtata atgggtcagc gacttacgtt cagtagcaag 600
cttaaccgta tagggcaggc gtaggaaagg agtccgaata gggcgttcag ttgctgggcg 660
tagacccgaa accaggtgat ctatccatgg ccaggatgaa ggtgcggtaa cacgtactgg 720
aggtccgaac ccactaacgt tgaaaagtta ggggatgagc tgtggatagg ggtgaaaggc 780
taaacaaacc tggaaatagc tggttctctc cgaaaactat ttaggtagtg cctcgtgtct 840
caccttcggg ggtagagcac tgtcatggtt ggggggtcta ttgcagatta ccccgccata 900
gcaaactccg aatatccgaag agtgcaatca cgggagacag acatcgggtg ctaacgtccg 960
gtgtcaagag ggaaacaacc cagaccgcca gctaaggtcc ccaaatatag ctaagtggga 1020
aacgaagtgg gaaggctaaa acagtcagga ggttggctta gaagcagcca ccctttaaag 1080
aaagcgtaat agctcactga tcgagtcgtc ctgcgcggaa gatgtaacgg ggctaagcta 1140
tatccgaag ctgcggatgc gtgctttgca cgatggtagg agagcgttcc gtaagcctgc 1200
gaaggtgcct tgtaaagggt gctggaggta tcggaagtgc gaatgctgac atgagtagcg 1260
ataaaggggg tgaaaggccc cctcgccgta agcccaaggt ttcctacgca acgttcatcg 1320
gcgtagggtg agtcggcccc taaggcgagg cagaaatgcg tagctgatgg gaagcaggtc 1380
aatattcctg caccattgtt agatgcgatg gggggacgga tcgcggaagg ttgtccgggt 1440
gttggaagtc ccggtcgctg cattggagaa ggcgcttagg caaatccggg cgcagaattc 1500
aagggtgtgg cgcgagctcc ttcgggagcg aagcaattgg aagtggttcc aagaaaagcc 1560
tctaagcttc agtctaacga tgaccgtacc gcaaaccgac acaggtgggc gagatgagta 1620
ttctaaggcg cttgagagaa ctcgggagaa ggaactcggc aaattggtac cgtaacttcg 1680
ggataaggta cgcccttgta gcttgactgg cctgcgccag gagggtgaag gggttgcaat 1740
aaactggtgg ctgcgactgt ttaataaaaa cacagcactc tgcaaacacg aaagtggacg 1800

tataggtgt gacgcctgcc cggtgccgga agattaaatg atggggtgca agctcttgat 1860
tgaagtcccg gtaaacggcg gccgtaacta taacggtcct aaggtagcga aattccttgt 1920
cgggtaagtt ccgacctgca cgaatggcgt aacgatggcc acactgtctc ctcccgagac 1980
tcagcgaagt tgaagtgttt gtgatgatgc aatctacccg cggctagacg gaaagacccc 2040
atgaacctt actgtagctt tgcatgtgac tttgaaccga tctgtgtagg ataggtggga 2100
ggctatgaaa ccggaacgct agtttcggtg gagccgtcct tgaaatacca ccctggttta 2160
tttgagttc taaccttggc ccgtgatccg ggtccgggac agtgcatggt aggcagtttg 2220
actgggcgg tctcctccca aagcgtaacg gaggagtacg aaggtacgct aggtacggtc 2280
ggaaatcgtg ctgatagtgc aatggcataa gcgtgcttaa ctgcgagacc gacaagtcga 2340
gcaggtgcga aagcaggtca tagtgatccg gtggttctgt atggaagggc catcgctcaa 2400
cggataaaag gtactctggg gataacaggc tgataccgcc caagagttca tatcgacggc 2460
ggtgtttggc acctcgatgt cggctcatct catcctgggg ctgtagccgg tcccaagggt 2520
atggctgttc gccatttaaa gaggtacgtg agctgggttt aaaacgtcgt gagacagttt 2580
ggtccctatc tgccgtgggc gttggatatt tgaagggggc tgctcctagt acgagaggac 2640
cggagtggac gaacctctgg tgtaccggtt gtcacgccag ggtcatcgcc gggtagctat 2700
gttcggaaga gataaccgct gaaagcatct aagcgggaaa ctcgccttaa gatgagatat 2760
ccctggggac tagatcccct tgaagggtcg ttcgagacca ggacgttgat aggtcaggtg 2820
tgtaagcgca gtaatgcgtt cagctaactg atactaattg cccgtaaggc ttgatcct 2878

<210> 67
<211> 2882
<212> DNA
<213> Burkholderia mallei

<400> 67
ggtcaagcga acaagtgcat gtggtggatg ccttggcgat cacaggcgat gaaggacgcg 60
gtagcctgcg aaaagctacg gggagctggc aaacgagctt tgatccgtag atgtccgaat 120
ggggaaaccc ggccctttg ggtcatccta gactgaatac ataggtctag tgaggcgaac 180
gcggtgaact gaaacatcta agtaaccgca ggaaaagaaa tcaaccgaga ttcccaaagt 240
agtggcgagc gaaatgggaa gagcctgtac tctttatttg tattgttagc cgaacgctct 300
ggaaagtgcg gccatagcag gtgatagccc tgtaggcgaa aacagtatga aagaactagg 360
tgtacgacaa gtagggcggg acacgtgaaa tcctgtctga agatgggggg accatcctcc 420
aaggctaaat actcgtgatc gaccgatagt gaaccagtac cgtgagggaa aggcgaaaag 480
aaccccggga ggggagtgaa atagatcctg aaaccgcatg catacaaaca gtcggagcct 540
cttcgggggt gacggcgtac cttttgtata atgggtcagc gacttacgtt cagtagcaag 600
cttaaccgaa tagggcaggc gtagcgaaag cgagtccgaa tagggcgttc agttgctgag 660
cgtagacccg aaaccaggtg atctatccat ggccaggatg aaggtgcggt aacacgtact 720
ggaggtccga acccactaac gttgaaaagt tagggggatga gctgtggata ggggtgaaag 780
gctaaacaaa cctggaaata gctggttctc tccgaaaact atttaggtag tgcctcgtgt 840
ctcaccttcg ggggtagagc actgtcatgg ttgggggggtc tattgcagat tacccgcgca 900
tagcaaactc cgaatacgga agagtgcaat cacgggagac agacatcggg tgctaacgtc 960
cggtgtcaag agggaaacaa cccagaccgc cagctaaggt ccccaaatat ggctaagtgg 1020
gaaacgaagt gggaaggcta aaacagtcag gaggttggct tagaagcagc cacccttaa 1080
agaaagcgta atagctcact gatcgagtcg tcctgcgcgg aagatgtaac ggggctaagc 1140
catataccga agctgcggat gcgagctagt ctcgcatggt aggagagcgt tccgtaagcc 1200
tgcgaaggtg cgttgaaaag cgtgctggag gtatccggaag tgcgaatgct gacatgagta 1260
gcgataaagg gggtgaaagg ccccctcgcc gtaagcccaa ggtttcctac gcaacgttca 1320
tcggcgtagg gtgagtcggc ccctaagggcg aggcagaaat gcgtagctga tgggaagcag 1380
gtcaatattc ctgcaccgtc gttagatgcg atgggggtgac ggatcgcgga aggttgtccg 1440
ggtgttggaa gtcccggtcg ctgcattgga gaaggcgctt aggcaaatcc gggcgcagga 1500
ttcaagggtg tggcgcgagc tccttcggga gcgaagcaat tggaagtggt tccaagaaaa 1560
gcctctaagc ttcagtctaa cgatgaccgt accgcaaacc gacacaggtg ggcgagatga 1620
gtattctaag gcgcttgaga gaactcggga gaaggaactc ggcaaattgg taccgtaact 1680
tcgggataag gtacgccctg tagcttgac tggcctgcgc cagaagggtg aaggggttgc 1740
aataaactgg tggctgcgac tgtttaataa aaacacagca ctctgcaaac acgaaagtgg 1800

acgtataggg tgtgacgcct gcccggtgcc ggaagattaa atgatggggt gcaagctctt 1860
gattgaagtc ccggtaaacg gcggccgtaa ctataacggt cctaaggtag cgaaattcct 1920
tgtcgggtaa gttccgacct gcacgaatgg cgtaacgatg gccacactgt ctcccccga 1980
gactcagcga agttgaagtg tttgtgatga tgcaatctac ccgcggctag acggaaagac 2040
cccatgaacc tttactgtag ctttgcattg gactttgaac cgatctgtgt aggatagctg 2100
ggaggctatg aaaccggaat gctagtttcg gtggagccgt ccttgaaata ccaccctggt 2160
ttgtttgagg tttctaacctt ggccgtgat ccgggtcggg gacagtgcat ggtaggcagt 2220
ttgactgggg cggtctcctc ccaaagcgta acggagggagt acgaaggtac gctaggtacg 2280
gtcggaaatc gtgctgatag tgcaatggca taagcgtgct taactgcgag accgacaagt 2340
cgagcaggtg cgaaagcagg tcatagtgat ccggtggttc tgtatggaag ggccatcgct 2400
caacggataa aaggtactct ggggataaca ggctgatacc gcccaagagt tcatatcgac 2460
ggcggtgttt ggcacctcga tgtcggctca tctcatcctg gggctgtagc cggtcccaag 2520
ggtatggctg ttcgccattt aaagaggtac gtgagctggg tttaaaacgt cgtgagacag 2580
tttggtccct atctgccgtg ggcgttggaa gtttgaaggg ggctgctcct agtacgagag 2640
gaccggagtg gacgaacctc tggtgtaccg gttgtgacgc cagtccgatc gccgggtagc 2700
tatgttcgga agagataacc gctgaaagca tctaagcggg aaactcgcct taagatgaga 2760
cttccccggg gacttgatcc ccttgaaggg tcgttcgaga ccaggacgtt gataggtcgg 2820
gtgtgtaagc gcagtaatgc gttcagctaa ccgatactaa ttgcccgtac ggcttgatcc 2880
ta 2882

<210> 68
<211> 2882
<212> DNA
<213> Burkholderia pseudomallei

<400> 68
ggtcaagcga acaagtgcat gtggtggatg ccttggcgat cacaggcgat gaaggacgcg 60
gtagcctgcg aaaagctacg gggagctggc aaacgagctt tgatccgtag atgtccgaat 120
ggggaaaccc ggccctttg ggtcatccta gactgaatac ataggtctag tgaggcgaac 180
gcggtgaact gaaacatcta agtaaccgca ggaaaagaaa tcaaccgaga ttcccaaagt 240
agtggcgagc gaaatgggaa gagcctgtac tctttatttg tattgttagc cgaacgctct 300
ggaaagtgcg gccatagcag gtgatagccc tgtaggcgaa aacagtatga aagaactagg 360
tgtacgacaa gtagggcggg acacgtgaaa tcctgtctga agatgggggg accatcctcc 420
aaggctaaat actcgtgatc gaccgatagt gaaccagtac cgtgagggaa aggcgaaaag 480
aaccccggga ggggagtgaa atagatcctg aaaccgcatg catacaaaca gtcggagcct 540
cttcgggggt gacggcgtac cttttgtata atggttcagc gacttacgtt cagtagcaag 600
cttaaccgaa tagggcaggc gtagcgaaag cgagtccgaa tagggcgttc agttgctggg 660
cgtagacccg aaaccaggtg atctatccat ggccaggatg aaggtgcggt aacacgtact 720
ggaggtccga acccactaac gttgaaaagt taggggatga gctgtggata ggggtgaaag 780
gctaaacaaa cctggaaata gctggttctc tccgaaaact atttaggtag tgcctcgtgt 840
ctcaccttcg ggggtagagc actgtcatgg ttggggggtc tattgcagat taccccgcca 900
tagcaaactc cgaataccga agagtgcaat cacgggagac agacatcggg tgctaacgtc 960
cggtgtcaag agggaaacaa cccagaccgc cagctaaggt ccccaaatat ggctaagtgg 1020
gaaacgaagt gggaaggcta aaacagtcag gaggttggct tagaagcagc cacccttaa 1080
agaaagcgta atagctcact gatcgagtcg tcctgcgcgg aagatgtaac ggggctaagc 1140
catataccga agctgcggat gcgagctagt ctcgcatggt aggagagcgt tccgtaagcc 1200
tgcgaaggtg cgttgaaaag cgtgctggag gtatccggaag tgcgaatgct gacatgagta 1260
gcgataaagg gggtgaaagg ccccctcgcc gtaagcccaa ggttcctac gcaacgttca 1320
tcggcgtagg gtgagtcggc ccctaaggcg aggcagaaat gcgtagctga tgggaagcag 1380
gtcaatattc ctgcaccgtc gttagatgcg atgggggcac ggatcgcgga aggttgtccg 1440
ggtgttggaa gtcccggtcg ctgcattgga gaaggcgctt aggcaaatcc gggcgcagga 1500
ttcaagggtg tggcgcgagc gctctagggc gcgaagcaat tggaagtggt tccaagaaaa 1560
gcctctaagc ttcagtctaa cgatgaccgt accgcaaacc gacacaggtg ggcgagatga 1620
gtattctaag gcgcttgaga gaactcggga gaaggaactc ggcaaattgg taccgtaact 1680
tcgggataag gtacgccctg gtagcttgac tggcctgcgc cagaagggtg aaggggttgc 1740

```
aataaactgg tggctgcgac tgtttaataa aaacacagca ctctgcaaac acgaaagtgg 1800
acgtatagag tgtgacgcct gcccggtgcc ggaagattaa atgatggggt gcaagctctt 1860
gattgaagtc ccggtaaacg gcggccgtaa ctataacggt cctaaggtag cgaaattcct 1920
tgtcgggtaa gttccgacct gcacgaatgg cgtaacgatg gccacactgt ctcctcccga 1980
gactcagcga agttgaagtg tttgtgatga tgcaatctac ccgcggctag acggaagaac 2040
cccatgaacc tttactgtag ctttgcattg gactttgaac cgatctgtgt aggataggtg 2100
ggaggctatg aaaccggaac gctagtttcg gtggagccgt ccttgaaata ccaccctggt 2160
ttgtttgagg ttctaacctt ggccgtgat ccgggtcggg gacagtgcat ggtaggcagt 2220
ttgactgggg cggtctcctc ccaaaggcgta acggaggagt acgaaggtac gctaggtacg 2280
gtcggaaatc gtgctgatag tgcaatggca taagcgtgct taactgcgag accgacaagt 2340
cgagcaggtg cgaaagcagg tcatagtgat ccggtggttc tgtatggaag ggccatcgct 2400
caacggataa aaggtactct ggggataaca ggctgatacc gcccaagagt tcatatcgac 2460
ggcggtgttt ggcacctcga tgtccggctca tctcatcctg gggctgtagc cggtcccaag 2520
ggtatggctg ttcgccattt aaagaggtac gtgagctggg tttaaaacgt cgtgagacag 2580
tttggtccct atctgccgtg ggcgttggaa gtttgaaggg ggctgctcct agtacgagag 2640
gaccggagtg gacgaacctc tggtgtaccg gttgtgacgc cagtcgcatc gccgggtagc 2700
tatgttcgga agagataacc gctgaaagca tctaagcggg aaactcgcct taagatgaga 2760
cttccccggg gacttgatcc ccttgaaggg tcgttcgaga ccaggacgtt gataggtcgg 2820
gtgtgtaagc gcagtaatgc gttcagctaa ccgatactaa ttgcccgtac ggcttgatcc 2880
ta                                                                2882
```

<210> 69
<211> 2890
<212> DNA
<213> Neisseria gonorrhoeae

<400> 69

```
ggtcaagtga ataagtgcat caggcggatg ccttggcgat gataggcgac gaaggacgtg 60
taagcctgcg aaaagcgcgg gggagctggc aataaagcta tgattccgcg atgtccgaat 120
ggggaaaccc actgcattct gtgcagtatc ctaagttgaa tacataggct tagagaagcg 180
aacccggaga actgaaccat ctaagtaccc ggaggaaaag aaatcaaccg agattccgca 240
agtagtggcg agcgaacgcg gaggagcctg tacgtaataa ctgtcgagat agaagaacaa 300
gctgggaagc ttgaccatag cgggtgacag tcccgtattc gaaatctcaa cagcggtact 360
aagccgtacga aaagtagggc gggacacgtg aaatcctgtc tgaatatggg gggaccatcc 420
tccaaggcta aatactcatc atcgaccgat agtgaaccag taccgtgagg gaaaggcgaa 480
aagaaccccg ggagggaagt gaaacagaac ctgaaacctg atgcatacaa acagtgggag 540
cgccctagtg gtgtgactgc gtacctttg tataatgggt caacgactta cattcagtag 600
cgagcttaac cggatagggg aggcgtaggg aaaccgagtc ttaatagggc gatgagttgc 660
tgggtgtaga cccgaaaccg agtgatctat ccatggtcag gttgaaggtg ccgtaacagg 720
tactggagga ccgaacccac gcatgttgca aaatgcgggg atgagctgtg ggtaggggtg 780
aaaggctaaa caaactcgga gatagctggt tctccccgaa aactatttag gtagtgcctc 840
gagcaagaca ctgatggggg taaagcactg ttatggctag gggggttattg caacttacca 900
acccatggca aactcagaat accatcaagt ggttcctcgg gagacagaca gcgggtgcta 960
acgtccgttg tcaagaggga aacaacccag accgccggct aaggtcccaa atgatagatt 1020
aagtggtaaa cgaagtggga aggcacagac agccaggatg ttggcttaga agcagccatc 1080
atttaaagaa agcgtaatag ctcactggtc gagtcgtcct gcgcggaaga tgtaacgggg 1140
ctcaaatcta taaccgaagc tgcggatgcc ggtttaccgg catggtaggg gagcgttctg 1200
taggctgatg aaggtgcatt gtaaagtgtg ctggaggtat cgaaagtgcg aatgttgaca 1260
tgagtagcga taaagcgggt gaaaagcccg ctcgccgaaa gcccaaggtt tcctacgcaa 1320
cgttcatcgg cgtagggtaa gtcggcccct aaggcgaggc agaaatgcgt agtcgatggg 1380
aaacaggtta atattcctgt acttgattca aatgcgatgt ggggacggag aaggttaggt 1440
tggcaagctg ttggaatagc ttgtttaagc cggtaggtgg aagacttagg caaatccggg 1500
ttttcttaac accgagaagt gatgacgagt gtctacggac acgaagcaac cgatacccag 1560
cttccaggaa aagccactaa gcttcagttt gaatcgaacc gtaccccaaa ccgacacagg 1620
tgggtaggat gagaattcta aggcgcttga gagaactcgg gagaaggaac tcggcaaatt 1680
```

gataccgtaa cttcgggaga aggtatgccc tctaaggtta aggacttgct ccgtaagccc 1740
cggagggtcg cagagaatag gtggctgcga ctgtttatta aaaacacagc actctgccaa 1800
cacgaaagtg gacgtatagg gtgtgacgcc tgccccggtgc cggaaggtta attgaagatg 1860
tgcaagcatc ggatcgaagc cccggtaaac ggcggccgta actataacgg tcctaaggta 1920
gcgaaattcc ttgtcgggta agttccgacc cgcacgaatg gcgtaacgat ggccacactg 1980
tctcctcccg agactcagcg aagttgaagt ggttgtgaag atgcaatcta cccgctgcta 2040
gacggaaaga ccccgtgaac ctttactgta ggttgcatt ggacttttgaa gtcacttgtg 2100
taggataggt gggaggcttg gaagcagaga cgccagtctc tgtggagtcg tccttgaaat 2160
accaccctgg tgtctttgag gttctaaccc agacccgtca tccgggtcgg ggaccgtgca 2220
tggtaggcag tttgactggg gcggtctcct cccaaagcgt aacggaggag ttcgaaggtt 2280
acctaggtcc ggtcggaaat cggactgata gtgcaatggc aaaaggtagc ttaactgcga 2340
gaccgacaag tcgggcaggt gcgaaagcag gacatagtga tccggtggtt ctgtatggaa 2400
gggccatcgc tcaacggata aaaggtactc cggggataac aggctgattc cgcccaagag 2460
ttcatatcga cggcggagtt tggcacctcg atgtcggctc atcacatcct gcggctgtag 2520
tcggtcccaa gggtatggct gttcgccatt taaagtggta cgtgagctgg gttttaaacg 2580
tcgtgagaca gtttggtccc tatctgcagt ggcgttggaa gtttgacggg gctgctccta 2640
gtacgagagg accggagtgg acgaacctct ggtgtaccgc ttgtaacgcc agttgcatag 2700
ccgggtagct aagttcggaa gagataagcg ctgaaagcat ctaagcgcga aactcgcctg 2760
aagatgagac ttccttgcg gtttaaccgc actaaagggt cgttcgagac caggacgttg 2820
ataggtgggg tgtggaagcg cggtaacgcg tgaagctaac ccatactaat tgcccgtgag 2880
gcttgactct 2890

<210> 70
<211> 2891
<212> DNA
<213> Neisseria meningitidis

<400> 70
gtcaagtgaa taagtgcatc aggtggatgc cttggcgatg ataggcgacg aaggacgtgt 60
aagcctgcga aaagcgcggg ggagctggca ataaagcaat gatccgcgca tgtccgaatg 120
gggaaaccca ctgcattctg tgcagtatcc taagttgaat acatagactt agagaagcga 180
acccggagaa ctgaaccatc taagtacccg gaggaaaaga aatcaaccga gattccgcaa 240
gtagtggcga gcgaacgcgg aggagcctgt acgtaataac tgtcgagata gaagaacaag 300
ctgggaagct tgaccatagt gggtgacagt cccgtattcg aaatctcaac agcggtacta 360
agcgtacgaa aagtagggcg gggcacgtga aatcctgtct gaatatgggg ggaccatcct 420
ccaaggctaa atactcatca tcgaccgata gtgaaccagt accgtgaggg aaaggcgaaa 480
agaaccccgg gaggggagtg aaacagaacc tgaaacctga tgcatacaaa cagtgggagc 540
gccctagtgg tgtgactgcg taccttttgt ataatgggtc aacgacttac attcagtagc 600
gagcttaacc gaatagggga ggcgtagggа aaccgagtct taatagggcg atgagttgct 660
gggtgtagac ccgaaaccga gtgatctatc catggccagg ttgaaggtgc cgtaacaggt 720
actggaggac cgaacccacg catgttgcaa aatgcgggga tgagctgtgg atagggggtga 780
aaggctaaac aaactcggag atagctggtt ctcccgaaa actatttagg tagtgcctcg 840
agcaagacac tgatgggggt aaagcactgt tatggctagg gggttattgc aacttaccaa 900
cccatggcaa actaagaata ccatcaagtg gttcctcggg agacagacag cgggtgctaa 960
cgtccgttgt caagagggaa acaacccaga ccgccagcta aggtcccaaa tgatagatta 1020
agtggtaaac gaagtgggaa ggcccagaca gccaggatgt tggcttagaa gcagccatca 1080
tttaaagaaa gcgtaatagc tcactggtcg agtcgtcctg cgcggaagat gtaacggggc 1140
tcaaatctat aaccgaagct gcggatgccg gtttaccggc atggtagggg agcgttctgt 1200
aggctgatga aggtgcattg taaagtgtgc tggaggtatc agaagtgcga atgttgacat 1260
gagtagcgat aaagcgggtg aaaagccccgc tcgccgaaag cccaaggttt cctgcgcaac 1320
gttcatcggc gtagggtgag tcggccccta aggcgaggca gaaatgcgta gtcgatggga 1380
aacaggttaa tattcctgta cttgattcaa atgcgatgtg gggacggaga aggttaggtt 1440
ggcaagctgt tggaatagct tgtttaagcc ggtaggtgga agacttaggc aaatccgggt 1500
cttcttaaca ccgagaagtg acgacgagtg tctacggaca cgaagcaacc gatacccacgc 1560
ttccaggaaa agccactaag cttcagtttg aatcgaaccg taccgcaaac cgacacaggt 1620

```
gggcaggatg agaatttctaa ggcgcttgag agaactcagg agaaggaact cggcaaattg 1680
ataccgtaac ttcgggagaa ggtatgccct ctaaggttaa ggacttgctc cgtaagcccc 1740
ggagggtcgc agagaatagg tggctgcgac tgtttattaa aaacacagca ctctgctaac 1800
acgaaagtgg acgtatagggg tgtgacgcct gcccggtgct ggaaggttaa ttgaagatgt 1860
gagagcatcg gatcgaagcc ccagtaaacg gcggccgtaa ctataacggt cctaaggtag 1920
cgaaattcct tgtcgggtaa gttccgaccc gcacgaatgg cgtaacgatg gccacactgt 1980
ctcctcctga gactcagcga agttgaagtg gttgtgaaga tgcaatctac ccgctgctag 2040
acggaaagac cccgtgaacc tttactgtag ctttgcattg gactttgaag tcacttgtgt 2100
aggataggtg ggaggcttag aagcagagac gccagtctct gtggagccgt ccttgaaata 2160
ccaccctggt gtctttgagg ttctaaccca gacccgtcat ccgggtcggg gaccgtgcat 2220
ggtaggcagt ttgactgggg cggtctcctc ccaaagcgta acggaggagt tcgaaggtta 2280
cctaggtccg gtcggaaatc ggactgatag tgcaatggca aaaggtagct taactgcgag 2340
accgacaagt cgagcaggtg cgaaagcagg acatagtgat ccggtggttc tgtatggaag 2400
ggccatcgct caacggataa aaggtactcc ggggataaca ggctgattcc gcccaagagt 2460
tcatatcgac ggcggagttt ggcacctcga tgtcggctca tcacatcctg gggctgtagt 2520
cggtcccaag ggtatggctg ttcgccattt aaagtggtac gtgagctggg tttaaaacgt 2580
cgtgagacag tttggtccct atctgcagtg ggcgttggaa gtttgacggg ggctgctcct 2640
agtacgagag gaccggagtg gacgaacctc tggtgtaccg gttgtaacgc cagttgcata 2700
gccgggtagc taagttcgga agagataagc gctgaaagca tctaagcgcg aaactcgcct 2760
gaagatgaga cttcccttgc ggtttaaccg cactaaagag tcgttcgaga ccaggacgtt 2820
gataggtggg gtgtggaagc gcggtaacgc gtgaagctaa cccatactaa ttgctcgtga 2880
ggcttgactc t                                                      2891
```

<210> 71
<211> 2891
<212> DNA
<213> Pseudomonas aeruginosa

```
<400> 71
ggtcaagtga agaagcgcat acggtggatg ccttggcagt cagaggcgat gaaagacgtg 60
gtagcctgcg aaaagcttcg gggagtcggc aaacagactt tgatccggag atctctgaat 120
gggggaaccc acctaggata actaggtat cttgtactga atccataggt gcaagaggcg 180
aaccaggggа actgaaacat ctaagtaccc tgaggaaaag aaatcaaccg agattccctt 240
agtagtggcg agcgaacggg gattagccct taagcttcat tgattttagc ggaacgctct 300
ggaaagtgcg gccatagtgg gtgatagccc cgtacgcgaa aggatctttg aagtgaaatc 360
gagtaggacg gagcacgaga aactttgtct gaacatgggg ggaccatcct ccaaggctaa 420
atactactga ctgaccgata gtgaaccagt accgtgaggg aaaggcgaaa agaaccccgg 480
agaggggagt gaaatagaac ctgaaaccgt atgcgtacaa gcagtgggag cctacttgtt 540
aggtgactgc gtacctttтg tataatgggt cagcgactta tattcagtgg caagcttaac 600
cgtataggct aggcgtagcg aaagcgagtc ttaataggggc gtttagtcgc tgggtataga 660
cccgaaaccg ggcgatctat ccatgagcag gttgaaggtt aggtaacact gactggagga 720
ccgaacccac tcccgttgaa aaggtagggg atgacttgtg gatcggagtg aaaggctaat 780
caagctcgga gatagctggt tctcctcgaa agctatttag gtagccgcctc atgtatcact 840
ctgggggtga gagcactgtt tcggctaggg ggtcatcccg acttaccaaa ccgatgcaaa 900
ctccgaatac ccagaagtgc cgagcatggg agacacacgg cgggtgctaa cgtccgtcgt 960
gaaaagggaa acaacccaga ccgccagcta aggtcccaaa gttgtggtta agtggtaaac 1020
gatgtgggaa ggcttagaca gctaggaggt tggcttagaa gcagccaccc tttaaagaaa 1080
gcgtaatagc tcactagtcg agtcggcctg cgcggaagat gtaacggggc tcaaaccaca 1140
caccgaagct gcgggtgtca cgtaagtgac gcggtagagg agcgttctgt aagcctgtga 1200
aggtgagttg agaagcttgc tggaggtatc agaagtgcga atgctgacat gagtaacgac 1260
aatgggtgtg aaaaacaccc acgccgaaag accaagggtt cctgcgcaac gttaatcgac 1320
gcagggttag tcggttccta aggcgaggct gaaaagcgta gtcgatggga aacaggttaa 1380
tattcctgta cttctggtta ctgcgatgga gggacggaga aggctaggcc agcttggcgt 1440
tggttgtcca agtttaaggt ggtaggctga aatcttaggt aaatccgggg tttcaaggcc 1500
gagagctgat gacgagtcgt cttttagatg acgaagtggt tgatgccatg cttccaagaa 1560
```

```
aagcttctaa gcttcaggta accaggaacc gtaccccaaa ccgacacagg tggtcgggta 1620
gagaatacca aggcgcttga gagaactcgg gtgaaggaac taggcaaaat ggcaccgtaa 1680
cttcgggaga aggtgcgccg gctagggtga aggatttact ccgtagctcg tggctggtcg 1740
aagataccag gccgctgcga ctgtttatta aaaacacagc actctgcaaa cacgaaagtg 1800
gacgtatagg gtgtgacgcc tgcccggtgc cggaaggtta attgatgggg ttagcgcaag 1860
cgaagctctt gatcgaagcc ccggtaaacg gcggccgtaa ctataacggt cctaaggtag 1920
cgaaattcct tgtcgggtaa gttccgacct gcacgaatgg cgtaacgatg gcggccgtgt 1980
ctccacccga gactcagtga aattgaaatc gctgtgaaga tgcagtgtat ccgcggctag 2040
acggaaagac cccgtgaacc tttactgtag ctttgcactg gactttgagc ctgcttgtgt 2100
aggataggtg ggaggctttg aagcgtggac gccagttcgc gtggagccat ccttgaaata 2160
ccaccctggc atgcttgagg ttctaacttct ggtccgtaat ccggatcgag gacagtgtat 2220
ggtgggcagt ttgactgggg cggtctcctc ctaaagagta acggaggagt acgaaggtgc 2280
gctcagaccg gtcggaaatc ggtcgcagag tataaaggca aaagcgcgct tgactgcgag 2340
acagacacgt cgagcaggta cgaaagtagg tcttagtgat ccggtggttc tgtatggaag 2400
ggccatcgct caacggataa aagggtactcc ggggataaca ggctgatacc gcccaagagt 2460
tcatatcgac ggcggtgttt ggcacctcga tgtcggctca tcacatcctg gggctgaagc 2520
cggtcccaag ggtatggctg ttcgccattt aaagtggtac gcgagctggg tttagaacgt 2580
cgtgagacag ttcggtccct atctgccgtg gacgtttgag atttgagagg ggctgctcct 2640
agtacgagag gaccggagtg gacgaacctc tggtgttccg gttgtcacgc cagtggcatt 2700
gccgggtagc tatgttcgga aaagataacc gctgaaagca tctaagcggg aaacttgcct 2760
caagatgaga tctcactggg aacttgattc ccctgaaggg ccgtcgaaga ctacgacgtt 2820
gataggctgg gtgtgtaagc gttgtgaggc gttgagctaa ccagtactaa ttgcccgtga 2880
ggcttgacca t                                                      2891
```

```
<210> 72
<211> 2886
<212> DNA
<213> Vibrio cholerae
```

```
<400> 72
ggttaagtga ctaagcgtac acggtggatg cctgggcagt cagaggcgat gaaggacgta 60
ctaacttgcg ataagcgcag ataaggcagt aagagccgtt tgagtctgcg atttccgaat 120
ggggaaaccc aactgcataa gcagttactg ttaactgaat acataggtta acagagcaaa 180
ccgggggaac tgaaacatct aagtaccccg aggagaagaa atcaaccgag attccggtag 240
tagcggcgag cgaacctgga ttagccctta agcactcggt gaagtaggtg aacaagctgg 300
aaagcttggc gatacagggt gatagccccg taaccgacgc ttcatcgagc gtgaaatcga 360
gtagggcggg acacgtgata tcctgtctga atatgggggg accatcctcc aaggctaaat 420
actcctgact gaccgatagt gaaccagtac cgtgaggaaa ggcgaaaaga acccctgtga 480
ggggagtgaa atagaacctg aaaccgtgta cgtacaagca gtaggagcac cttcgtggtg 540
tgactgcgta ccttttgtat aatgggtcag cgacttatat tcagtggcaa ggttaaccgt 600
ataggggagc cgtagcgaaa gcgagtctta actgggcgct cagtctctgg atatagaccc 660
gaaaccgggt gatctagcca tgggcaggtt gaaggttgag taacatcaac tggaggaccg 720
aaccgactaa tgttgaaaaa ttagcggatg acttgtggct aggggtgaaa ggccaatcaa 780
actcggagat agctggttct ccccgaaagc tatttaggta gcgcctcgga cgaatactac 840
tgggggtaga gcactgttaa ggctaggggg tcatcccgac ttaccaaccc tttgcaaact 900
ccgaatacca gtaagtacta tccgggagac acacggcggg tgctaacgtc cgtcgtggag 960
agggaaacaa cccagaccgc cagctaaggt cccaaagtat tgctaagtgg gaaacgatgt 1020
gggaaggctc agacagctag gatgttggct tagaagcagc catcatttaa agaaagcgta 1080
atagctcact agtcgagtcg gcctgcgcgg aagatgtaac ggggctaagc aatacaccga 1140
agctgcggca atatctttta gatattgggt aggggagcgt tctgtaagcc gttgaaggtg 1200
aatcgtaagg tttgctggag gtatcagaag tgcgaatgct gacatgagta acgacaaagg 1260
gggtgaaaaa cctcctcgcc ggaagaccaa gggttcctgt ccaacgttaa tcggggcagg 1320
gtgagtcgac ccctaaggtg aggccgaaag gcgtaatcga tgggaaacgg gttaatattc 1380
ccgtacttct gactattgcg atgggggtac ggagaaggct agtggggcca ggcgacggtt 1440
gtcctggttc aagtgcgtag gcttgagagt taggtaaatc cggctctctc taaggctgag 1500
```

acacgacgtc gagctactac ggtagtgaag tcattgatgc catgcttcca ggaaaagcct 1560
ctaagcttca gatagtcagg aatcgtaccc caaacegaca caggtggtcg ggtagagaat 1620
accaaggcgc ttgagagaac tcgggtgaag gaactaggca aaatggtacc gtaacttcgg 1680
gagaaggtac gctcttgatg gtgaagtccc tcgcggatgg agctgacgag agtcgcagat 1740
accaggtggc tgcaactgtt tattaaaaac acagcactgt gcaaaatcgc aagatgacgt 1800
atacggtgtg acgcctgccc ggtgccggaa ggttaattga tggggttagc gcaagcgaag 1860
ctcttgatcg aagccccggt aaacggcggc cgtaactata acggtcctaa ggtagcgaaa 1920
ttccttgtcg ggtaagttcc gacctgcacg aatggcgtaa tgatggccac gctgtctcca 1980
cccgagactc agtgaaattg aaatcgctgt gaagatgcag tgtacccgcg gctagacgga 2040
aagaccccgt gaacctttac tacagcttgg cactgaacat tgaacctaca tgtgtaggat 2100
aggtgggagg ctatgaagac gtgacgccag ttgcgttgga gccgtccttg aaataccacc 2160
cttgtatgtt tgatgttcta acttagaccc gttatccggg ttgaggacag tgcctggtgg 2220
gtagtttgac tggggcggtc tcctcccaaa gagtaacgga ggagcacgaa ggtgggctaa 2280
tcacggttgg acatcgtgag gttagtgcaa tggcataagc ccgcttaact gcgagaatga 2340
cggttcgagc aggtgcgaaa gcaggtcata gtgatccggt ggttctgtat ggaaggtgca 2400
tcgctcaacg gataaaaggt actccgggga taacaggctg ataccgccca agagttcata 2460
tcgacggcgg tgtttggcac ctcgatgtcg gctcatcaca tcctggggct gaagtcggtc 2520
ccaagggtat ggctgttcgc catttaaagt ggtacgcgag ctgggtttag aacgtcgtga 2580
gacagttcgg tccctatctg ccgtgggcgt tggaagattg aaggggctg ctcctagtac 2640
gagaggaccg gagtggacga acctctggtg ttcgggttgt gtcgccagac gcattgcccg 2700
gtagctaagt tcggaattga taagcgctga aagcatctaa gcgcgaagcg agccctgaga 2760
tgagtcttcc ctgacagttt aactgtccta aagggttgtt cgagactaga acgttgatag 2820
gcagggtgtg taagcgttgt gaggcgttga gctaacctgt actaattgcc cgtgaggctt 2880
aaccat 2886

<210> 73
<211> 2906
<212> DNA
<213> Yersinia enterocolitica

<220>
<221> modified_base
<222> (1168)..(1178)
<223> N = A, C, G or T/U

<400> 73
ggttaagcga ccaagcgtac acggtggatg cctaggcagt cagaggcgat gaaggacgtg 60
Ctaatctgcg aaaagcgtcg gtaaggtgat atgaaccgtt ataaccgacg ataccccgaat 120
ggggaaaccc agtgcaattc gttgcactat tgcatggtga atacatagcc atgcaaggcg 180
aaccgggggga actgaaacat ctaagtaccc cgaggaaaag aaatcaaccg agattccccc 240
agtagcggcg agcgaacggg gaggagccca gaaacctgaat cagcgtatgt gttagtggaa 300
gcgtctggaa agtcgcacgg tacagggtga tagtcccgta cacaaaaatg catatgttgt 360
gagttcgatg agtagggcgg gacacgtgac atcctgtctg aatatggggg gaccatcctc 420
caaggctaaa tactcctgac tgaccgatag tgaaccagta ccgtgaggga aaggcgaaaa 480
gaaccccggc gaggggagtg aaacagaacc tgaaaccgtg tacgtacaag cagtgggagc 540
accttcgtgg tgtgactgcg tacctttttgt ataatgggtc agcgacttat attttgtagc 600
aaggttaacc gaatagggga gccgtaggga aaccgagtct taactgggcg aatagttgca 660
aggtatagac ccgaaacccg gtgatctagc catgggcagg ttgaaggttg ggtaacacta 720
actggaggac cgaaccgact aatgttgaaa aattagcgga tgacttgtgg ctgggggtga 780
aaggccaatc aaaccgggag atagctggtt ctccccgaaa gctatttagg tagcgcctcg 840
tgaactcatc ttcgggggta gagcactgtt tcggctaggg ggtcatcccg acttaccaaa 900
ccgatgcaaa ctccgaatac cgaagaatgt tatcacggga gacacacggc gggtgctaac 960
gtccgtcgtg aagagggaaa caaccccagac cgccagctaa ggtcccaaag tcatggttaa 1020
gtgggaaacg atgtgggaag gcacagacag ccaggatgtt ggcttagaag cagccatcat 1080
ttaaagaaag cgtaatagct cactggtcga gtcggcctgc gcggaagatg taacggggct 1140

aaaccatgca ccgaagctgc ggcagcgnnn nnnnnnnnnn nnnnnnnngg ggagcgttct 1200
gtaagccgtt gaaggtgacc tgtgaggctt gctggaggta tcagaagtgc gaatgctgac 1260
ataagtaacg ataatgcggg tgaaaaaccc gcacgccgga agaccaaggg ttcctgtcca 1320
acgttaatcg gggcagggtg agtcgacccc taaggcgagg ctgaaaggcg tagtcgatgg 1380
gaaacaggtt aatattcctg tacttggtgt tactgcgaag gggggacgga gaaggctatg 1440
ctagccgggc gacggttgtc ccggtttaag catgtaggcg gagtgaccag gtaaatccgg 1500
ttgcttatca acgctgaggt gtgatgacga gtcactacgg tgatgaagta gttgatgcca 1560
tgcttccagg aaaagcctct aagcatcagg taacatgaaa tcgtacccca aaccgacaca 1620
ggtggtcagg tagagaatac tcaggcgctt gagagaactc gggtgaagga actaggcaaa 1680
atggtgccgt aacttcggga gaaggcacgc tgacacgtag gtgaagcggt ttaccccgtgg 1740
agctgaagtc agtcgaagat accagctggc tgcaactgtt tattaaaaac acagcactgt 1800
gcaaacacga aagtggacgt atacggtgtg acgcctgccc ggtgctggaa ggttaattga 1860
tggggtcagc gcaagcgaag ctcttgatcg aagccccggt aaacggcggc cgtaactata 1920
acggtcctaa ggtagcgaaa ttccttgtcg ggtaagttcc gacctgcacg aatggcgtaa 1980
tgatggccag gctgtctcca cccgagactc agtgaaattg aactcgctgt gaagatgcag 2040
gtacccgcg gcaagacgga aagaccccgt gaacctttac tatagcttga cactgaacat 2100
tgagccttga tgtgtaggat aggtgggagg catagaagtg tggacgccag tctgcatgga 2160
gccaaccttg aaataccacc ctttaatgtt tgatgttcta actcggcccc gtaatccggg 2220
gtgaggacag tgtcaggtgg gtagtttgac tggggcggtc tcctcccaaa gagtaacgga 2280
ggagcacgaa ggttagctaa tcacggtcgg acatcgtgag gttagtgcaa aggcataagc 2340
tagcttcact gcgagagtga cggctcgagc aggtacgaaa gtaggtctta gtgatccggt 2400
ggttctgaat ggaagggcca tcgctcaacg gataaaaggt actccgggga taacaggctg 2460
ataccgccca agagttcata tcgacggcgg tgttggcac ctcgatgtcg gctcatcaca 2520
tcctggggct gaagtaggtc ccaagggtat ggctgttcgc catttaaagt ggtacgcgag 2580
ctgggtttag aacgtcgtga gacagttcgg tccctatctg ccgtgggcgy tggarraytg 2640
agrggggctg ctcctagtac gagaggaccg gagtggacgm atcactggtg ttcgggttgt 2700
catgccaatg gcaytgcccg gtagctaaat kcggaagaga taasygctga aagcatctaa 2760
gcrsgaaact tgccycgaga tgagtttctcc ctgagactac aagtctcctg aaggaacgtt 2820
gaagacgacg acgttgatag gcyyggtgtg taagcgcgag ttggcgttga gctaaccggt 2880
actaatgaac cgtgaggctt aaccttt 2906

<210> 74
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 74
gggttgcgct cgttacggga ctt 23

<210> 75
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 75
gggttgcgct cgttgccgga ctt 23

<210> 76
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 76
tccccactgc tgcctcccgt agg 23

<210> 77
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 77
caacatctca cgacacgagc tga 23

<210> 78
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 78
tccccactgc tgcctcccgt agg 23

<210> 79
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 79
ttaccgcggc tgctggcacg ga 22

<210> 80
<211> 23

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 80
ccccgtcaat tcctttgagt ttc 23

<210> 81
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 81
caacatctca cgacacgagc tga 23

<210> 82
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 82
tttcaccttt ccctcacggt act 23

<210> 83
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 83
ggttcttttc acctttccct cgc 23

<210> 84
<211> 23
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic Primer

<400> 84
tggtttcagg ttctatttca ctc 23

<210> 85
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 85
tttaaccgac aaggaatttc gc 22

<210> 86
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 86
ggttcttttc acctttccct cgc 23

<210> 87
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 87
taacctggtc gtaac 15

<210> 88
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 88

ccccccccc cccc 14

<210> 89
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 89
gcccctaacc tcgtcg 16

<210> 90
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 90
cggccctagc cgggtcgtac ctccgg 26

<210> 91
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 91
cggccctaac ctggtcgtaa ctccgg 26

<210> 92
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 92
aggcttcgat cccgggatcc gcg 23